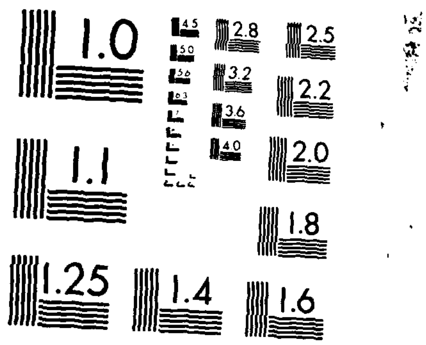
AD-A163 359 COMPONENTS OF VERBAL INTELLIGENCE(U) YALE UNIV NEW HAVEN CT DEPT OF PSYCHOLOGY R J STERNBERG 30 DEC 85 N00014-83-K-0013 1/1

UNCLASSIFIED F/G 5/10 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A163 359

Final Report:

Components of Verbal Intelligence

Contract N0001483K0013 from the Office of Naval Research and the Army Research Institute

October, 1982 - September, 1985

Robert J. Sternberg
Yale University

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, and by the U.S. Army Research Institute, Basic Research, under Contract No. N0001483K0013, Contract Authority Identification Number, NR 154-505.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

AD-A163 359

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): Final Report 12-85	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Yale University, Department of Psychology; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Office of Naval Research and the U.S. Army Research Institute
6c. ADDRESS (City, State, and ZIP Code): New Haven, Connecticut 06520	7b. ADDRESS (City, State, and ZIP Code): Arlington, Virginia 22217; and Alexandria, Virginia 22333
8a. NAME OF FUNDING/SPONSORING ORGANIZATION; 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N0001483K0013
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
		NR154-505	

11. TITLE (Include Security Classification)

"Components of Verbal Intelligence" (Unclassified)

12. PERSONAL AUTHOR(S)

Robert J. Sternberg

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 10/1/82 TO 9/30/85	1985 December 30	90

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Decontextualization, information processing, causal inferenc· learning from context, novelty, nonentrenched concepts, verbal comprehension, vocabulary acquisition

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This document constitutes a final report for Contract N0001483K0013 from the Office of Naval Research and the Army Research Institute.

The purpose of this project was to develop and test a theory of the components of verbal intelligence. Alternative theoretical frameworks for understanding verbal intelligence are reviewed, and then a componential theory of verbal comprehension is proposed. The theory specifies the information-processing components, context cues, and mediating variables underlying acquisition of word meanings from context. A number of experiments testing and supporting the theory are described, including experiments involving both internal and external context. Instructional experiments are also described, and it is concluded that the theory is well supported by the data, and moreover, that it can serve as a useful basis for training people in how to learn meanings of words from context. The theory is extended to novel kinds of concepts as well, and it is shown that the learning of novel concepts involves an interaction between linguistic and conceptual unfamiliarity.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL: Susan Chipman	22b. TELEPHONE (Include Area Code): (202)696-4318; 22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

19. ABSTRACT

In a series of experiments on causal inference involving verbal versus symbolic-abstract materials, it is shown that although a core of logical operations is applied to both verbal and abstract materials, special processes are involved when people reason about meaningful verbal materials. Although the research reported focuses primarily upon top-down kinds of verbal processing, a final experiment shows that there is a continuum of levels of processing along which bottom-up and top-down tasks can be classified, and that as a result, a final theory of verbal comprehension will have to account for processing along this entire continuum, rather than for top-down and bottom-up elements as separate entities.

Final Report:

Components of Verbal Intelligence

Contract N0001483K0013 from the Office of Naval Research and the Army Research Institute

October, 1982 - September, 1985

Robert J. Sternberg
Yale University

Final Report: Components of Verbal Intelligence

Verbal comprehension refers to a person's ability to understand linguistic materials, such as newspapers, magazines, textbooks, lectures, and the like. Verbal comprehension has been recognized as an integral part of intelligence in both psychometric theories (e.g., Guilford, 1967; Thurstone, 1938; Vernon, 1971) and information-processing theories (e.g., Carroll, 1976; Hunt, 1978; Sternberg, 1980) and has, under a variety of aliases, been an important topic of research in both differential and experimental psychology.

The theoretical construct of verbal comprehension can be and has been operationalized in a number of different ways. Most often, it is directly measured by tests of vocabulary, reading comprehension, and general information. Indeed, vocabulary has been recognized not only as an excellent measure of verbal comprehension, but also as one of the best single indicators of a person's overall level of intelligence (e.g., Jensen, 1980; Matarazzo, 1972). The importance of verbal comprehension in general, and of vocabulary in particular, to the measurement of verbal intelligence is shown by the fact that both of the two major individual scales of intelligence--the Stanford-Binet and the Wechsler--contain vocabulary items, and by the fact that many group tests also contain vocabulary items (which may be presented in any of a number of forms, e.g., synonyms, antonyms, verbal analogies with very low frequency terms, and so on). Because of its importance both in the theory and measurement of intelligence and in everyday interactions with the environment, it seems important to understand the antecedents of observable individual differences in vocabulary levels.

The theory of verbal comprehension proposed in this chapter comprises two parts. The first part is a theory of how an aspect of verbal comprehension--learning from context--develops. The second part is a theory of information processing in verbal comprehension, that is, of the skills one uses in one's current verbal functioning. Thus, the first theory accounts for how "crystallized" ability becomes crystallized; the second theory accounts for how crystallized ability is utilized in information processing. Each of these two subtheories of verbal comprehension as a whole will be considered in turn, with acquisition considered in consecutive sections of the chapter. Consider first the question of acquisition.

In order for my presentation of the theory of the acquisition of verbal comprehension skills to be fully meaningful, it must first be placed in the context of other efforts toward the same or similar goals. There have been three major approaches to understanding the origins and development of verbal comprehension. These three major approaches are considered briefly below.

Alternative Cognitive Approaches to the Acquisition of Verbal Comprehension Skills

The three major approaches to the acquisition of verbal comprehension skills are a "knowledge-based" approach, a "bottom-up" approach, and a "top-down" approach. The knowledge-based approach deals with the role of prior information in the acquisition of new information. The bottom-up approach deals with speed of execution of certain very basic mechanistic cognitive processes. The top-down approach deals with higher-order utilization of cues in complex verbal materials. Consider each of these three approaches, and some of the research that has been done under each.

The Knowledge-Based Approach

The knowledge-based approach assigns a central role to old knowledge in the acquisition of new knowledge. Although "knowledge" is often referred to in the sense of domain-specific information, the knowledge-based approach can also

encompass research focusing on general world knowledge, knowledge of structures or classes of text (as in story grammars), and knowledge about strategies for knowledge acquisition and application (see, e.g., Bisanz & Voss, 1981). Proponents of this approach differ in the respective roles they assign to knowledge and process in the acquisition of new knowledge. A fairly strong version of the approach is taken by Keil (1984), who argues for the primacy of knowledge over process in cognitive development.

Proponents of the knowledge-based approach usually cite instances of differences between expert and novice performance--in verbal and other domains--that seem more to derive from knowledge differences than from processing differences. For example, Keil (1984) suggests that development in the use of metaphor and in the use of defining features of words seems to be due more to differential knowledge states than to differential use of processes or speed of process execution. Chi (1978) has shown that whether children's recall performance is better than that of adults depends upon the knowledge domain in which the recall takes place, and particularly, upon the relative expertise of the children and adults in the respective domains. Chiesi, Spilich, and Voss (1979) and Spilich, Vesonder, Chiesi, and Voss (1979) have shown the importance of prior knowledge about baseball in the acquisition of new information about this topic. In related research, Chase and Simon (1973) found that differences between expert and novice performance in chess seemed largely to be due to differential knowledge structures rather than processes (but see Charness, 1981).

A rather extensive study of the relations between verbal comprehension ability and word knowledge was conducted by Curtis (1981). Curtis assessed levels of word knowledge for each of 37 college students for each of 24 stem words and 72 answer options. She did her analysis in terms of three levels of word knowledge: (a) decoding, measured by the percentage of words that each subject was able to read aloud correctly, (b) semantic familiarity, which was defined as the percentage of words for which a subject, in addition to correct decoding, could produce any accurate semantic information at all, regardless of the nature of that information, and (c) semantic precision, which was the percentage of responses for which the subject provided a correct decoding and a synonym or correct explanation of the word's meaning. Scores from a 115-item multiple-choice vocabulary test (which drew from several standardized tests) were regressed on each of the three components of word knowledge plus speed of lexical access. Curtis found that overall, the vocabulary test seemed most to measure word familiarity. In a related analysis, Curtis regressed Verbal Scholastic Aptitude Test scores on measures of familiarity, precision, and speed. Overall, precision had the highest weight and speed the lowest (and nonsignificant) weight. Interestingly, however, there was an apparent interaction between skill level of the subjects and the obtained regression weights. For high verbal subjects, precision had a slightly higher standardized regression weight than did familiarity, whereas for lower verbal subjects, familiarity had a much higher regression weight than did precision. Speed had a trivial weight in each case. These results suggest that for a difficult verbal ability test such as the SAT, different skills may be measured at different points along the ability continuum.

I have no argument with the position that the knowledge base is highly important in understanding differences in current performance between experts and novices in both verbal and nonverbal domains. But accounts that slight the role of information processing in the development of expertise seem to beg an important question, namely, that of how the differences in knowledge states came about in the first place. For example, why did some people acquire better vocabularies than others? Or in the well-studied domain of chess, why is it that of two individuals given equally intensive and extensive exposure to the game, one will acquire the knowledge structures needed for expertise, and the other will not? In sum, I accept the importance of old knowledge in the acquisition of new knowledge. But I do not believe the overemphasis on process that characterized some previous research should be replaced by an overemphasis on knowledge in present research. Rather, it should be recognized that knowledge and process work interactively in complex ways. What is needed is to understand what these ways are.

The Bottom-Up Approach

Bottom-up research has emerged from the tradition of investigation initiated by Earl Hunt (e.g., Hunt, 1978, 1980; Hunt, Lunneborg, & Lewis, 1975) and has been followed up by a number of other investigators (e.g., Jackson & McClelland, 1979; Keating & Bobbitt, 1978; see also Perfetti & Lesgold, 1977, for a related approach). According to Hunt (1978), two types of processes underlie verbal comprehension ability--knowledge-based processes and mechanistic (information-free) processes; Hunt's approach has emphasized the latter kind of process. Hunt et al. (1975) studied three aspects of what they called "current information processing" that they believed to be key determinants of individual differences in developed verbal ability. These were:

> (a) sensitivity of overlearned codes to arousal by incoming stimulus information, (b) the accuracy with which temporal tags can be assigned, and hence order information can be processed, and (c) the speed with which the internal representations in STM and intermediate term memory (ITM, memory for events occurring over minutes) can be created, integrated, and altered. (p. 197)

The basic hypothesis motivating this work is that individuals varying in verbal ability differ even in these low-level mechanistic skills--skills that are free from any contribution of disparate knowledge or experience. Intelligence tests are hypothesized to measure indirectly these basic information-processing skills by measuring directly the products of these skills, both in terms of their past contribution to the acquisition and storage of knowledge (such as vocabulary), and their present contribution in the current processing of information.

Perfetti (1983) has suggested that four basic, bottom-up verbal processes underlie some, but not all, individual differences in reading ability. These processes are (a) word decoding (i.e., the transformation of a printed input into one or more of its corresponding linguistic forms), (b) letter recognition (i.e., recognizing the constituent letters of a word,) (c) name retrieval (i.e., accessing the location in memory and producing the name of a verbal string), and (d) semantic access (i.e., activation of the meaning components of a word stored in memory). In a series of experiments, Perfetti has demonstrated that each of these skills is related to measures of verbal ability and reading skill (e.g., Perfetti & Hogaboam, 1975).

In a typical experiment employing the bottom-up approach to verbal comprehension, subjects are presented with the Posner and Mitchell (1967) letter-matching task. The task comprises two experimental conditions, a physical-match condition and a name-match condition. In the physical-match condition, subjects are presented with pairs of letters that either are or are not physical matches (e.g., "AA" or "bb" versus "Aa" or "Ba"). In the name-match condition, subjects are presented with pairs of letters that either are or are not name matches (e.g., "Aa," "BB," or "bB" versus "Ab," "ba," or "bA"). Subjects must identify the letter pair either as a physical match (or mismatch) or as a name match (or mismatch) as

rapidly as possible. The typical finding in these experiments is that the difference between mean name match and physical match times within a group of subjects correlates about -.3 with scores on a test of verbal ability.

The finding described above seems to be widely replicable, but its interpretation is a matter of dispute (Carroll, 1981; Hogaboam & Pellegrino, 1978; Sternberg, 1981). I have been and remain concerned that .3-level correlations are abundant in both the abilities and personality literatures (indeed, they are rather low as ability correlations go), and provide a relatively weak basis for causal inference. A further concern is that most of the studies that have been done on the name minus physical match difference have not used adequate discriminant validation procedures. When such procedures are used, and perceptual speed is considered as well as verbal ability, this difference seems to be much more strongly related to perceptual speed than it is to verbal ability (Lansman, Donaldson, Hunt, & Yantis, 1982; Cornelius, Willis, Blow, & Baltes, 1983), although these findings are subject to alternative interpretations. Thus, the obtained correlation with verbal ability may reflect, at least in part, variance shared with perceptual abilities of the kind that the letter-matching task would seem more likely to measure. But whatever may be the case here, it seems likely that speed of lexical access plays <u>some</u> role in verbal comprehension, and what remains to be clarified is just what this role is.

The role of speed of lexical access seems to increase as the complexity of the decision motivating the access increases. Goldberg, Schwartz, and Stewart (1977) had subjects perform a comparison task at three levels of complexity. The first level was simply physical comparison (e.g., are "A" and "A" a physical match?). The second level was homophone comparison (e.g., are "here" and "hear" a sound match?). The third level was taxonomic-category comparison (e.g., are "cat" and "dog" a match with respect to belonging in the category of animals?). These authors found a greater difference between the performance of pre-identified high- and low-verbal subjects on the homophone and taxonomic comparison tasks than they did on the physical comparison task. The results suggest that the more "top-down" the level of comparison, the greater the difference between high and low verbals.

A similar effect of "level of processing" can be seen in the studies that have been done relating performance on a sentence-picture comparison task to verbal ability (Baddeley, 1968; Hunt, Lunneborg, & Lewis, 1975; Lansman, Donaldson, Hunt, & Yantis, 1982). In this paradigm, initiated by Clark and Chase (1972), subjects are shown a sentence, such as "Plus is above star," and a picture that either correctly represents the information in the sentence or that does not correctly represent this information. Subjects are required to indicate as quickly as possible whether the sentence and picture do or do not correspond. Correlations between latencies on this task and scores on verbal ability tests are typically in the -.4 to -.6 range. These correlations are higher than those obtained for the Posner and Mitchell (1967) task (see Hunt, 1984; Hunt et al., 1975), presumably because the complexity of processing required is greater. These results lead quite naturally to a consideration of the last, "top-down," approach to understanding verbal comprehension skills.

<u>Top-Down Approach</u>

Top-down processing refers to expectation- or inference-driven processing, or to "knowledge-based" processing, to use Hunt's (1978) terminology. Top-down processing has been an extremely popular focus for research in the past decade, with many researchers attempting to identify and predict the sorts of inferences a person is likely to draw from a text and how these inferences (or lack thereof) will

affect text comprehension (see, e.g., Kintsch & van Dijk, 1978; Rieger, 1975; Rumelhart, 1980; Schank & Abelson, 1977). Usually, top-down researchers look at how people combine information actually present in the text with their own store of world knowledge to create a new whole representing the meaning of the text (e.g., Bransford, Barclay, & Franks, 1972).

The first of a small handful of investigators who looked at the use of inference in the acquisition of word meanings from context were Werner and Kaplan (1952), who proposed that learning from context provides a major source of vocabulary development. They devised a task in which subjects were presented with an imaginary word followed by six sentences using that word. The subjects' task was to guess the meaning of the word on the basis of the contextual cues they were given. They found that performance improves gradually with age, although the various processes underlying performance did not necessarily change gradually. They did not, however, provide an explicit model of what these processes are.

Daalen-Kapteijns and Elshout-Mohr (1981) pursued the Werner-Kaplan approach by having subjects think aloud while solving Werner-Kaplan type problems. They found, among other things, that high- and low-verbal subjects learn word meanings differently, with high verbals performing a deeper analysis of the possibilities for a new word's meaning than was performed by low-verbal subjects. In particular, the high-verbal subjects used a well-formulated strategy for figuring out word meaning, whereas the low-verbal subjects seemed not to.

Keil (1981) presented children in grades kindergarten, 2, and 4 with simple stories in which an unfamiliar word was described by a single paragraph. An example of such a story is "THROSTLES are great, except when they have to be fixed. And they have to be fixed very often. But it's usually very easy to fix throstles." Subjects were asked what else they knew about the new word (here, "throstle"), and what sorts of things the new word described. Keil found that even the youngest children could make sensible inferences about the general categories denoted by the new terms and about the properties the terms might reasonably have (see also Keil, 1979, 1981, 1984).

Jensen (1980) has suggested that vocabulary is an excellent measure of intelligence "because the acquisition of word meanings is highly dependent on the deduction of meaning from the contexts in which the words are encountered" (p. 146). Marshalek (1981) has tested this hypothesis using a faceted vocabulary test, although he did not directly measure learning from context. He found that subjects with low reasoning ability did, in fact, have major difficulties inferring meanings of words. Moreover, reasoning was related to vocabulary measures at the lower end of the vocabulary difficulty distribution but not at the higher end. Together, these findings suggest that a certain level of reasoning ability may be prerequisite for extraction of word meaning. Above this level, the importance of reasoning begins to decrease rapidly.

Quite a different top-down approach to the understanding of verbal comprehension has been taken by Daneman (1984; Daneman & Carpenter, 1980). The theory motivating this research is that functional working memory capacity--involving complex processing in the working memory store--is critically involved in individual differences in reading skill. Daneman's paradigm requires subjects to read a successive string of unrelated sentences, and to memorize the last word in each sentence. For example, one set of three sentences used was

> (1) He had patronized her when she was a schoolgirl and teased her when she was a student. (2) He had an odd elongated skull which sat on his shoulders like a pear on a dish. (3) The products of digital electronics will play an important role in your future.

Scores on this span measure are surprisingly low, usually in the range of 3-5 words. More interestingly, however, is the fact that scores on the span measure typically correlate in the .4 to .6 range with scores on tests of verbal ability. Even higher correlations have been obtained with scores on measures of ability to integrate linguistic information. It thus appears that Daneman's span measure provides a promising route to understanding the role of working memory in reading.

Summary. To summarize, I have described three basic approaches to understanding the cognitive bases of verbal comprehension. These approaches--a knowledge-based one, a bottom-up one, and a top-down one--are complementary, and ultimately would all have to be incorporated to understand fully the nature and development of verbal comprehension. In the next section, I present our own approach to understanding the antecedents and development of verbal comprehension skills, which is largely top-down in character.

Theory of Learning from Context

Our theory of the development of verbal skills emphasizes learning from context (see Sternberg & Powell, 1983; Sternberg, Powell, & Kaye, 1983). We believe that the ability to infer the meanings of unfamiliar words from context deserves a prominent place within a discussion of verbal comprehension for three reasons. First, a theory describing how people use context to infer the meanings of words could tell us much about vocabulary building skills. Identifying what types of information people of different ability levels use to construct a tentative definition of a word and how additional information influences a working definition of a word could also tell us about how to train vocabulary acquisition skills. Second, a theory of learning from context can help explain why vocabulary is the single best predictor of verbal intelligence. Our hypothesis is that learning from context reflects important vocabulary acquisition skills, the net products of which could, in theory, be measured by the full extent of one's vocabulary. Thus, according to our view, vocabulary tests are such good predictors of one's overall verbal intelligence because they reflect one's ability to acquire new information. Third, a theory of learning from context is useful in illuminating the relationship between the more fluid, inferential aspects of verbal intelligence, usually measured by tests of verbal analogies, and the more crystallized, knowledge-based aspects of verbal intelligence, usually measured by vocabulary tests (see Horn & Cattell, 1966). Learning from context thus provides a way of integrating the two aspects of verbal ability--comprehension and vocabulary--and of placing vocabulary acquisition within the framework of general cognitive theories of language comprehension.

Two basic ideas underlie our theory of learning from context. The first idea pertains to why some verbal concepts are easier to learn than others: The difficulty of learning a new verbal concept is in large part a function of the degree of facilitation (or inhibition) of learning provided by the context in which the new verbal concept is embedded; the same or very similar contextual elements that facilitate (or inhibit) learning of the concept are hypothesized also to facilitate (or inhibit) later retrieval of the concept and also its transfer to new situations. The second idea pertains to why some individuals are better at learning verbal concepts than are others: Individual differences in verbal comprehension can be traced in large part to differences in people's abilities to use context to learn verbal concepts, and to be wary of contextual elements that inhibit learning; the same or very similar sources of individual differences are hypothesized to be involved in people's differential abilities later to retrieve verbal concepts and to transfer these concepts appropriately to new situations.

The theory distinguishes between those aspects of vocabulary acquisition that lie strictly outside the individual, that is, contextual cues present in the verbal

context that convey various types of information about the word, and those aspects of vocabulary acquisition that lie at least partially within the individual, that is, mediating variables that affect the perceived usefulness of the contextual cues. The contextual cues determine the quality of a definition that theoretically can be inferred for a word from a given context. The mediating variables specify those constraints imposed by the relationship between the previously unknown word and the context in which the word occurs that affect how well a given set of cues will be actually utilized, by an individual, in a particular task and situation. Moreover, the theory specifies the processes by which the cues and mediating variables are utilized. These various aspects of the theory will now be explained in turn.

Theory of Decoding of External Context

During the course of one's reading (or other encounters with words), one commonly comes upon words whose meanings are unfamiliar. When such words are encountered, one may attempt to utilize the external context in which the words occur in order to figure out the meanings of the words. Our theory specifies contextual cues, mediating variables, and processes that influence the likelihood that these meanings will be correctly inferred.

Contextual cues. Context cues are hints contained in a passage that facilitate (or, *in theory and sometimes in practice, impede*) *deciphering* the meaning of an unknown word. We propose that context cues can be classified into eight categories, depending upon the kind of information they provide. These context cues include the following:

1. Temporal cues. Cues regarding the duration or frequency of X (the unknown word), or regarding when X can occur; alternatively, cues describing X as a temporal property (such as duration or frequency) of some Y (usually a known word in the passage). (Example: I saw the WEX last night.)

2. Spatial cues. Cues regarding the general or specific location of X, or possible locations in which X can sometimes be found; alternatively, cues describing X as a spatial property (such as general or specific location) of some Y. (Example: I saw the WEX in the forest.)

3. Value cues. Cues regarding the worth or desirability of X, or regarding the kinds of affects X arouses; alternatively, cues describing X as a value (such as worth or desirability) of some Y. (Example: I was afraid of the WEX.)

4. Stative descriptive cues. Cues regarding properties of X (such as size, shape, color, odor, feel, etc.); alternatively, cues describing X as a stative descriptive property (e.g., shape or color) of some Y. (Example: The WEX was gray.)

5. Functional descriptive cues. Cues regarding possible purposes of X, actions X can perform, or potential uses of X; alternatively, cues describing X as a possible purpose, action, or use of Y. (Example: The WEX snarled at me.)

6. Causal/Enablement cues. Cues regarding possible causes of or enabling conditions for X; alternatively, cues describing X as a possible cause or enabling condition for Y. (Example: The WEX made me run for my life!)

7. Class membership cues. Cues regarding one or more classes to which X belongs, or other members of one or more classes of which X is a member; alternatively, cues describing X as a class of which Y is a member. (Example: The WEX is a canine.)

8. Equivalence cues. Cues regarding the meaning of X, or contrasts (such as antonymy) to the meaning of X; alternatively, cues describing X as the meaning (or a contrast in meaning) of some Y. (Example: The WEX I saw, like most other wolves, was fierce.)

An example of the use of some of these cues in textual analysis *might help*

concretize our descriptive framework. Consider the sentence, "At dawn, the BLEN arose on the horizon and shone brightly." This sentence contains several external contextual cues that could facilitate one's inferring that BLEN probably means SUN. "At dawn" provides a temporal cue, describing when the arising of the BLEN occurred; "arose" provides a functional descriptive cue, describing an action that a BLEN could perform; "on the horizon" provides a spatial cue, describing where the arising of the BLEN took place; "shone" provides another functional descriptive cue, describing a second action a BLEN could do; finally, "brightly" provides a stative descriptive cue, describing a property (brightness) of the shining of the BLEN. With all these different cues, it is no wonder that most people would find it very easy to figure out that the neologism BLEN is a synonym for the familiar word, SUN.

We make no claim that the categories we have suggested are mutually exclusive, exhaustive, or independent in their functioning. Nor do we claim that they in any sense represent a "true" categorization scheme of context cues. We have found, however, that this classification scheme is useful in understanding subjects' strategies in deriving meanings of words from context. Not every type of cue will be present in every context, and even when a given cue is present, our theory proposes that the usefulness of the cue will be mediated by the sorts of variables to be described in the next section.

<u>Mediating variables</u>. Whereas the contextual cues describe the types of information that might be used to infer the meaning of a word from a given verbal context, they do not at all address the problems of recognition of the applicability of a description to a given concept, weeding out irrelevant information, or integration of the information gleaned into a coherent model of the word's meaning. For this reason, a set of mediating variables is also proposed that specifies relations between a previously unknown word and the passage in which it occurs, and that mediates the usefulness of the contextual cues. Thus, whereas the contextual cues specify the particular kinds of information that might be available for an individual to use to figure out the meanings of unfamiliar words, the mediating variables listed below specify those variables that can affect, either positively or negatively, the application of the contextual cues present in a given situation.

1. <u>Number of occurrences of the unknown word</u>. A given kind of cue may be absent or of little use in a given occurrence of a previously unknown word, but may be present or of considerable use in another occurrence. Multiple occurrences of an unknown word increase the number of available cues and can actually increase the usefulness of individual cues if readers integrate information obtained from cues surrounding the multiple occurrences of the word. For example, the meaning of a given temporal cue may be enhanced by a spatial cue associated with a subsequent appearance of the unknown word, or the temporal cue may gain in usefulness if it appears more than once in conjunction with the unknown word. On the other hand, multiple occurrences of an unfamiliar word can also be detrimental if the reader has difficulty integrating the information gained from cues surrounding separate appearances of the word, or if only peripheral features of the word are reinforced and are therefore incorrectly interpreted as being of central importance to the meaning of the unfamiliar word.

2. <u>Variability of contexts in which multiple occurrences of the unknown word appear</u>. Different types of contexts, for example, different kinds of subject matter or different writing styles, and even just different contexts of a given type, such as two different illustrations within a given text of how a word can be used, are likely to supply different types of information about the unknown word.

Variability of contexts increases the likelihood that a wide range of types of cues will be supplied about a given word, and thus increases the probability that a reader will get a full picture of the scope of a given word's meaning. In contrast, mere repetition of a given unknown word in essentially the same context as that in which it previously appeared is unlikely to be as helpful as a variable-context repetition, because few or no really new cues are provided regarding the word's meaning. Variability can also present a problem in some situations and for some individuals: If the information is presented in a way that makes it difficult to integrate across appearances of the word, or if a given individual has difficulties in making such integrations, then the variable repetitions may actually obfuscate rather than clarify the word's meaning. In some situations and for some individuals, a stimulus overload may occur, resulting in reduced rather than increased understanding.

3. <u>Importance of the unknown word to understanding the context in which it is embedded</u>. If a given unknown word is judged to be necessary for understanding the surrounding material in which it is embedded, the reader's incentive for figuring out the word's meaning is increased. If the word is judged to be unimportant to understanding what one is reading (or hearing), one is unlikely to invest any great effort in figuring out what the word means. Whereas in explicit vocabulary-training situations, the individual may always be motivated to infer a word's meanings, in real-world situations, this will not be the case. Thus, a question of interest from the perspective of our model is the extent to which an individual reader can recognize which words are important to a passage, and which are not. In some cases, it really may not be worth the individual's time to figure out a given word's meaning. It is possible to distinguish between importance at different levels of text organization. We distinguish between the sentence and paragraph levels, that is, the importance of a given word to understanding the meaning of the sentence in which it occurs, and to understanding the meaning of the paragraph in which it occurs. The ability to recognize the importance of a word to understanding context may be seen as one form of comprehension monitoring of the form studied by Markman (1977, 1979), Flavell (1981), Collins and Smith (1982), and others.

4. <u>Helpfulness of surrounding context in understanding the meaning of the unknown word</u>. A given cue can be differentially helpful depending upon the nature of the word whose meaning is to be inferred and upon the location of the cue in the text relative to the word whose meaning is to be inferred. Consider first an example of how the nature of the word can affect cue helpfulness. A temporal cue describing when a <u>diurnal</u> event occurs would probably be more helpful than a spatial cue describing where the event occurs in aiding an individual to figure out that "diurnal" means <u>daily</u>. In contrast, a spatial cue would probably be more helpful than a temporal cue in figuring out that <u>ing</u> is a low-lying pasture. It is unrealistic to expect a given kind of cue to be equally helpful in figuring out the meanings of all kinds of words. Consider now an example of how the location of the cue in the text relative to the word whose meaning is to be inferred can affect cue helpfulness. If a given cue occurs in close proximity to the word whose meaning is unknown, then there is probably a relatively high likelihood that the cue will be recognized as relevant to inferring the unknown word's meaning. If the cue is separated from the known word by a substantial portion of text, the relevance of the cue may never be recognized; indeed, the cue may be misinterpreted as relevant to an unknown word to which it is more proximal. The helpfulness of context cues may also be mediated by whether the cue comes before or after the unknown word. Rubin (1976), for example, found that context occurring before the placement of a blank in a cloze test was more helpful to figuring out what word

should go in the blank than was context occurring after the placement of the blank.

5. Density of unknown words. If a reader is confronted with a high density of previously unknown words, he or she may be overwhelmed and be unwilling or unable to use available cues to best advantage. When the density of unknown words is high, relatively more text is occupied by unknown and therefore unhelpful words (for figuring out meanings of other words), and it can be difficult to discern which of the cues that are available apply to which of the words that are unknown. In such a situation, utilization of a given cue may depend upon figuring out the meaning of some other unknown word, in which case the usefulness of that cue (and very likely of other cues as well) is decreased.

6. Concreteness of the unknown word and the surrounding context. Concrete concepts are generally easier to apprehend, in part because they have a simpler meaning structure. Familiar concrete concepts such as tree, chair, and pencil are relatively easy to define in ways that would satisfy most people; familiar abstract concepts such as truth, love, and justice, however, are extremely difficult to define in ways that would satisfy large numbers of people. Indeed, each of these abstract concepts has been the subject of multiple books, none of which have provided "definitive" definitions. Moreover, the ease of inferring the meaning of the word will depend upon the concreteness of the surrounding description. A concrete concept such as ing might appear more opaque embedded in a passage about the nature of reality than it would embedded in a passage about the nature of food sources; similarly, an abstract concept such as pulchritude (beauty) might be more easily apprehended in a passage about fashion models than in one about eternal versus ephemeral qualities.

7. Usefulness of previously known information in cue utilization. Inevitably, the usefulness of a cue will depend upon the extent to which past knowledge can be brought to bear upon the cue and its relation to the unknown word. The usefulness of prior information will depend in large part upon a given individual's ability to retrieve the information, to recognize its relevance, and then to apply it appropriately.

Knowledge-acquisition components and representation of information. The theory of external decontextualization also relies upon three knowledge-acquisition components, or processes.

1. Selective encoding. Selective encoding involves sifting out relevant from irrelevant information. When new words are presented in actual contexts, cues relevant to decontextualization are embedded within large amounts of irrelevant information. A critical task facing the individual is that of sifting out the "wheat from the chaff": recognizing just what information in the passage is relevant for word decontextualization.

2. Selective combination. Selective combination involves combining selectively encoded information in such a way as to form an integrated, plausible definition of the previously unknown word. Simply sifting out the relevant cues is not enough to arrive at a tentative definition of the word: One must know how to combine the cues into an integrated knowledge representation.

3. Selective comparison. Selective comparison involves relating newly acquired information to information acquired in the past. Deciding what information to encode and how to combine it does not occur in a vacuum. Rather, encoding and combination of new knowledge is guided by retrieval of old information. A cue will be all but useless if it cannot somehow be related to past knowledge.

Verbal information is theorized to be represented in terms of a network-type model that is similar in some respects to the node models found in Rumelhart and

Norman's (1975) and Collins and Loftus's (1975) models of semantic memory. A given concept is represented as the "center" of a network describing the concept. Nodes emanate from the concept describing its properties. Nodes for different concepts are connected via the concept names, which serve as the origin for nodes with descriptive attributes. Unlike in other network models, the kinds of nodes extending from the concept, and from other nodes, correspond to the properties of cues used to understand word meanings, as specified by the proposed theory of cue utilization. For example, spatial cues are fed into (where?) nodes, functional-descriptive cues are fed into (do?) nodes, stative descriptive cues are fed into (look like?) nodes, class membership cues are fed into (what?) nodes, equality cues are fed into (equals?) nodes, and so on. Each node has associated with it both an attribute (e.g., an attribute for (look like?) might be "gray") and an identification of the attribute as being necessary, sufficient, or characteristic of the concept. An example of this form of representation for "kangaroo" is shown in Figure 1.

(INSERT FIGURE 1 ABOUT HERE)

How is the proposed representation developed during the course of acquisition of a verbal concept? This question is addressable in terms of the processes, cues, and moderating variables in the proposed theory. A general description of the development of representations will be presented first, followed by a specific example of how this development occurs. It is assumed that initial processing is done sentence-by-sentence, although further processing may follow if a subject reviews a passage and uses higher-order units (e.g., pairs of sentences) as a basis for further understanding. A person begins building up a representation of text as soon as he or she starts reading the text.

The subject begins selective encoding of information about the to-be-defined (target) word from the first sentence in which the word appears. The target word at this point becomes the center point of a new network; characteristic and defining attributes can "grow" into appropriate nodes in this network, which is constructed in working memory. This information also activates matching information stored in long-term-memory networks. This activated knowledge then influences what further facts in the passage will be selectively encoded and fed into the nodes in the newly forming network. As more information about the new word is selectively encoded and incorporated into the new network, the subject's activated knowledge base in long-term memory is reduced: Concepts that might have helped define the word are now found to be excluded by the additional information, and hence can be dropped from active consideration. When the subject has finished processing all of the information in the passage, he or she will select concepts from the long-term memory knowledge base whose nodes are still activated. The full network structures for these concepts is then compared to the newly formed network structure. If no such concepts exist (e.g., all concepts in long-term memory have been excluded as possible meanings of the word), the subject will either reprocess the passage, or else view the new network structure as corresponding to a new concept nonidentical to any already in long-term memory. If one or more such concepts exist, the subject will compare defining attributes of the target word's network to defining attributes of the networks for each of the possible meanings. The subject will then select the activated concept that has the most defining attributes in common with the target concept, create a new concept, or else seek further information. A new concept will be based upon an extension of that concept in long-term memory that is closest to the new one, appropriately modified so as to take into account those nodes in the new representation that do not match the nodes of the old representation. So, for example, if ing is found to be closest in its representation to pasture, but to differ from pasture in having

nodes describing the ing as low-lying, then ing will be defined as a "pasture that is low-lying." In some instances, the given information may allow the individual to propose a definition that he or she knows is at a more general level than the correct meaning of the word. There was simply insufficient information fully to restrict the meaning of the new concept.

Errors in understanding the meaning of a new word can occur in at least three ways. First, information about the new word's meaning as provided in the text will inevitably be incomplete. Thus, one may not have sufficient basis for choosing among alternative possible meanings stored in long-term memory, or one may provide a new but incomplete definition. Second, information about the new word's meaning may be misencoded. A cue in the passage may be misconstrued, so that the representation one builds up is simply wrong. Third, information in the passage may be properly encoded, but lead to an incorrect representation of the new word because the information is misleading. In such a case, cues may actually serve to lead a subject astray.

The ideas expressed above can be made more explicit and concrete by considering an example of how the representation of a word is constructed. An illustration of the development of a representation of a word is shown in Figure 2. It is important to note that the buildup shown is for a given hypothetical individual: There will be individual differences, and probably major ones, in the representational buildups of various people as a function of their decontextualization skills (application of knowledge-acquisition components to contextual cues as mediated by the mediating variables) and prior knowledge. The subject is shown the following brief story about a BLUMEN and is asked to figure out what a BLUMEN is.

> He first saw a BLUMEN during a trip to Australia. He had just arrived from a business trip to India, and felt very tired. Looking out at the plain, he saw a BLUMEN hop across it. It was a typical marsupial, getting its food by chewing on the surrounding plants. Squinting because of the bright sunlight and an impending headache, he noticed a young BLUMEN securely fastened in an opening in front of its mother.

(INSERT FIGURE 2 ABOUT HERE)

In Step 1 (see Figure 2), the subject considers the first sentence in the passage, and selectively encodes two facts, that the individual saw a BLUMEN and that he first saw it on a trip to Australia. The first cue, a stative-descriptive cue, indicates that the BLUMEN is visible; the second cue, a spatial-locative cue, indicates that BLUMENS can be found in Australia. In selective combination, the representation of BLUMEN grows two nodes, a (look-like?) node for the stative-descriptive cue and a (where?) node for the spatial cue. In selective comparison, the subject's knowledge about things that can be seen in Australia is activated in long-term memory. The names (network central entries) of these concepts are placed into working memory and the subject constructs a list corresponding to possible meanings of BLUMEN. This list will be reduced in successive steps as entries in LTM and even classes of entries are found to be irrelevant to the new word's meaning. As each entry is deleted from the list of possible meanings in working memory, the nodes in long-term memory corresponding to that entry are deactivated.

In Step 2, the subject considers the second sentence. Because the subject is now using his or her activated knowledge to guide what should be selectively encoded, none of the information in this sentence is perceived as relevant to the task at hand (figuring out the meaning of BLUMEN). The reason for this is that the new information is uninformative (again, for this individual) with respect to

BLUMENS, and their visibility in Australia. No further encoding, combination, or comparison is done.

In Step 3, the subject considers the third sentence, selectively encoding that BLUMENS can be found on plains (a spatial cue related to the subject's knowledge that plains exist in Australia) and that BLUMENS hop (a functional-descriptive cue related to the subject's knowledge about what some animals do in Australia). The subject now selectively combines the new information with the information already in the BLUMEN network, adding two new nodes. The subject grows "on plains" out of "Australia," so that, according to the modified network, BLUMENS are now found on plains in Australia. "Hops" is fed into the (do?) node. Because of this newly encoded and combined information, the subject can eliminate some of the names of concepts that he or she is holding in working memory. In particular, by selectively comparing the new information with the activated information in long-term memory corresponding to the names of concepts being considered as possible word meanings in working memory, the subject can eliminate all names that do not represent objects that hop or that are found on plains.

In Step 4, the subject selectively encodes the facts that a BLUMEN is a marsupial and that it chews plants. (For someone who didn't know what a marsupial is, the information might be ignored, or an attempt might be made to infer the meaning of marsupial from context.) In selective combination, two further nodes are grown, a further (do?) node and a (what?) node. The two new attributes that have been added to the network can be used by the subject to reduce further the number of concept names he or she is holding in working memory. In particular, the subject can now eliminate names of objects whose network representations in long-term memory (which are still activated) do not represent marsupials that chew plants. Thus, selective comparison continues to reduce the relevant prior knowledge base at the same time that selective encoding and combination are increasing the relevant new knowledge base.

In Step 5, the subject fails selectively to encode any of the information in the sentence as relevant to the meaning of BLUMEN. This failure may derive either from his or her not realizing that there is relevant information, or from his or her not knowing how to use the given information. This failure illustrates how some of the moderating variables specified by the proposed theory can affect the build-up of a word representation. With successive presentations of the word, the chances are improved that the representation will be more nearly complete; one presentation often will not be enough to achieve anything approaching a complete representation.

The subject now checks whether there are any concept names left in working memory that meet all of the constraints of the representation he has built up. If there is only one such concept, the subject compares defining attributes of the network representation corresponding to the already stored word. If the attributes match or pass a criterion for being close enough, the subject defines the new word in terms of the old, in this case, "kangaroo." If the attributes do not match or are not close enough to accept the old name as a definition, the subject either offers a definition that represents a new concept different from any already stored in long-term memory, or else goes back to the passage and tries to obtain further information. If multiple old concept names are left in working memory, the subject compares the defining attributes of the new concept to the defining attributes of all of the remaining old concepts, and selects the best of the options if it is good enough (over criterion); if it is not good enough (i.e., it is under criterion), the subject either defines a wholly new concept or else goes back to the passage for further information. Again, this new concept will be a modification of

the best old-fitting concept, with the modification reflecting the mismatch between the new concept and the old one.

In conclusion, definitions of new words are constructed by adding defining and characteristic attributes on to new network representations at the same time that one reduces in size a list of possible meanings for the new word. The reduction is accomplished by comparing attributes of the new word to attributes of the listed words (as stored in long-term memory), and removing from the list words whose attributes do not match the attributes of the new word. Eventually, one is left with a built-up representation of the new word, and a usually reduced list of possible meanings. One then compares in working memory attributes of the new word to attributes of each of the words on the reduced list, and either (a) chooses one as the correct meaning if the match is close enough, (b) comes to view the new word as a new concept because it does not match any old concepts in memory, or (c) returns to the passage for more information. Should one see the word again in another context, one can return to the building-up process, and use the new information to refine and elaborate the network representation of the new word. This refinement and elaboration is more likely to lead to a correct definition in the final comparison process whereby defining attributes of new and old words are compared.

<u>Data testing the theory of external decontextualization</u>. We have some preliminary data regarding the validity of the proposed theory. In particular, we have tested only the cue-utilization and moderating-variable subtheories (Sternberg & Neuse, 1983; Sternberg & Powell, 1983).

<u>Experiment 1</u>. The theory was first tested (Sternberg & Powell, 1983) by asking 123 high school students to read 32 passages of roughly 125 words in length that contained embedded within them from 1 to 4 extremely low-frequency words. Thirty-seven of the words (all nouns) were used in the passages; each target word could appear from 1 to 4 times, resulting in 71 presentations altogether. Passages were equally divided among four different writing styles: literary, newspaper, scientific, and historical. An additional sample passage was written in the literary style. Consider it here as an example of the kinds of passages used:

> Two ill-dressed people--the one a tired woman of middle years and the other a tense young man--sat around a fire where the common meal was almost ready. The mother, Tanith, peered at her son through the <u>oam</u> of the bubbling stew. It had been a long time since his last <u>ceilidh</u> and Tobar had changed greatly; where once he had seemed all legs and clumsy joints, he now was well-formed and in control of his hard, young body. As they ate, Tobar told of his past year, re-creating for Tanith how he had wandered long and far in his quest to gain the skills he would need to be permitted to rejoin the company. Then all too soon, their brief <u>ceilidh</u> over, Tobar walked over to touch his mother's arm and quickly left.

The students' task was to define, as best they could, each of the low-frequency words within each passage (except for multiple occurrences of a single word within a given passage, which required only a single definition). Students were not permitted to look back to earlier passages and definitions in making their current responses.

Qualities of definitions were rated independently by three trained raters. Because mean inter-rater reliability was .92, an average of the three ratings was used as a definition-goodness score for each word for each subject. These averages were then averaged over subjects to obtain a mean goodness-of-definition rating for each word. The main independent variables were ratings of the number or strength of the occurrences of our contextual cues and moderating variables (with

the exact nature of the rating depending upon the independent variable) with respect to their roles in helping in the deciphering of the meaning of each low-frequency word in the passages.

Theory testing was done via multiple regression. We used a stepwise multiple-regression procedure in which we allowed only three variables plus a regression intercept to enter into our final models. The decision to limit the number of variables was made on the basis of our judgment of the degree of refinement of our data, and in the hope of minimizing the risks of capitalization upon chance that inhere in stepwise regression. Because of multicollinearity (correlation among) independent variables, it was not possible to make strong inferences regarding the "true" subsets of variables that were differentially relevant from one passage style to the next. Variables that entered into at least one of four regressions were enablement, stative-descriptive, functional-descriptive, and equivalence cues, plus moderating variables of helpfulness and importance. The correlations between predicted and observed goodness ratings were .92 for literary passages, .74 for newspaper passages, .85 for science passages, and .77 for history passages. All of these values were statistically significant.

We concluded on the basis of these data that the contextual cues and moderating variables proposed by our subtheories provided good prediction of the goodness-of-definition data, although we certainly do not believe that our model accounted for all of the reliable variance. Indeed, the square roots of the internal-consistency reliability coefficients (based on all possible split halves of subjects) for our four data sets, which place upper limits on the values of R, were all .98 or above, showing that there was a considerable amount of reliable variance not accounted for by the fitted model. Nevertheless, the fits of the model subsets seemed sufficiently high to merit some optimism regarding our initial attempts to understand differential word difficulty in learning from context. Moreover, performance on the task was successful in distinguishing high from low verbal subjects: Definition goodness ratings for individual subjects correlated .62 with IQ, .56 with vocabulary, and .65 with reading comprehension scores. The data, although extremely limited, are consistent with the notion that the proposed theory of cognitive competence is on the right track, at least in the domain of verbal declarative knowledge.

This first study had some clear limitations. Independent variables were nonorthogonal (multicollinear), resulting in difficulties isolating the effects of each variable; the possibility of interactions among model variables was not examined; and the population was limited to upper-middle class high school students. A second study was designed to expand upon the first by removing these limitations.

<u>Experiment 2</u>. A second study, also done in collaboration with Janet Powell, was designed to circumvent certain of the limitations of the first study. Subjects were 190 students in the ninth through twelfth grades of a suburban high school. Each subject received 13 short passages, each containing one extremely low-frequency noun. In one condition, the words were defined for the subjects. In another condition, the words were not defined. The subjects' task was to rate the helpfulness of each of a series of segments for defining the unknown word. Ratings were made on a 1 (low) to 7 (high) scale. The following passage fragments show two segmentation conditions: one with longer segments, and one with shorter segments:

> ...and when he removed his hat,//she, who preferred "ageless" men,// eyed his increasing <u>phalacrosis</u>//and grimaced.

...and when he removed//his hat//she,//who preferred//"ageless" men,//eyed//his increasing//phalacrosis//and grimaced.

Thus, subjects would have to rate the helpfulness of each segement--long or short--to defining phalacrosis (baldness). A set of raters independently rated each segment for what context cues, if any, it contained. For example, "and grimaced," a short segment, provides a value cue, and "when he removed his hat" provides a temporal and a spatial cue. It should be noted that for the most part, even segments containing cues contained only weak cues. Only a few segments contained strong cues. The mean helpfulness ratings were then averaged across segments for each cue type, in order to compute a mean helpfulness for each cue type. This averaging procedure yielded helpfulness weights for the various context cues, without introducing the multicollinearity problems that emerge from multiple regression.

The mean importances were 1.81 for setting (temporal and spatial) cues, 2.37 for value cues, 2.72 for stative-descriptive cues, 2.42 for functional cues, 2.51 for causal cues, 2.71 for class membership cues, and 3.20 for equivalence cues (including antonyms). The mean rating when there was no cue at all was 1.61. Because each of these means is based upon roughly 1000 observations, on the average, the differences in means are highly reliable. As would be expected, equivalence cues were the most helpful. These were followed by stative-descriptive and class membership cues (both of which deal with static properties) and then by functional and causal cues (both of which deal with active properties). Value cues were less helpful, and setting cues were the least helpful of all. It thus appears that all of the cues help somewhat, relative to the no-cue control segments, but they help differentially.

In this experiment, the cues were very weak, and no attempt was made to balance cues precisely. A followup experiment was done with more salient and more closely controlled cues. This experiment also tested aspects of the theory other than the context cues, and incorporated elements of training in the design.

Experiment 3. In a third experiment (Sternberg & Neuse, 1983), we tested 81 sophomores and juniors in an inner-city high school. The subjects were divided into two basic groups, a training group (59 subjects) and a control (no-training) group (22 subjects). The mean IQ of the subjects was 97, with a standard deviation of 11.

The experimental design in the third experiment involved seven independent variables: (a) training group (experimental, control), (b) testing time (pretest, posttest), (c) test format (blank, nonword), (d) clue type (stative descriptive, functional descriptive, class membership), (e) unknown word type (abstract, concrete), (f) restrictiveness of context with respect to the meaning of the unknown word (low, high), and (g) sentence function of the unknown word (subject, predicate). These variables were completely crossed with respect to each other. Treatment group was a between-subjects variable; all other variables were within subject and were manipulated via a faceted testing arrangement. Two different test forms were used, and half the subjects received the first form as a pretest and the second form as a posttest; the other half of the subjects received the reverse arrangement. Test items, involving either neologisms or blanks (cloze procedure), were each presented in the context of a single sentence. Subjects were asked either to define the neologism or to fill in the blank, as appropriate. There were 48 items on each test. Scores on the pretest were correlated .74 with an IQ test (Henmon-Nelson) given before training, and .71 with an alternative form of the test

given after training. Scores on the posttest were correlated .65 and .64, respectively, with the two administrations of the IQ test.

The training sequence was spread out over six sessions. The topics covered were (a) What is context? (b) Using sentence context, (c) 20 questions (spotting clue types), (d) Cues I (temporal, spatial, stative-descriptive, equivalence), (e) Using paraphrase to figure out word meanings, (f) Cues II (functional-descriptive, causal), and (g) Mystery words (neologisms presented in sentences or paragraphs). The six class periods proved ample to cover this range of theory-based material.

In the experimental group, significant main effects were obtained for testing time (posttest higher than pretest), clue type (stative-descriptive hardest, functional-descriptive in-between, category membership easiest), context restrictiveness (higher restrictive more difficult than lower restrictive), and sentence function (predicates harder than subjects). In the control group, significant main effects were obtained for clue type (same ordering of means as above) and restrictiveness of context (same ordering of means as above). Thus, there was a significant pre- to posttest gain in the trained group, but not the untrained group. However, the interaction between group and training effect was not statistically significant. In addition, there were a number of statistically significant interactions between independent variables, suggesting that model effects were not wholly independent and additive, but rather were interactive with each other.

Taken as a whole, these results suggest (a) that subsets of the cues and moderating variables do have additive effects that can be quantified and isolated, (b) that the additive effects are supplemented by interactive ones, and (c) that at least some training of decontextualization skills is possible. The set of results is thus supportive of the ideas in the theory of verbal decontextualization, but emphasizes the need to consider interactions as well as main effects in analyses of model fits.

The greatest disappointment in this experiment was the weakness of the training effects. I believed that enough had been learned from this experiment that it would now be possible to design an experiment that explicitly looked at training effects, without the "distraction" of also testing other things. The next experiment was therefore directed specifically at obtaining improvements in decontextualization ability through theoretically based training.

Experiment 4. In this experiment, 150 New Haven area adults (nonstudents) of roughly average intelligence were divided into one of five conditions. There were three training conditions and two control conditions. Subjects in all three training conditions and one of the control conditions received exactly the same practice words and passages, but differed in the instruction they received (if any) regarding the passages. Passages were similar to those in Experiment 1.

The 30 subjects in each of these conditions were given a 25-item pretest and a 25-item posttest measuring skill in figuring out word meanings, as well as other tests. The pretest and posttest were transfer tests, in that they measured skill in figuring out word meanings: They did not merely test recall of words in the practice materials. Our goal was not to train specific vocabulary, but rather to train vocabulary-learning skills. All words in the experiment were extremely rare English-language words. The same pretest and posttest words were used in each condition, and training words were the same across conditions. Items were scored on a 0-2 point scale, for a maximum score of 50 points per test. Each training session lasted 45 minutes, exclusive of the various kinds of testing, which brought session length to 2½ hours. The conditions, which were between-subjects, were as

follows:

1. Process training. Subjects were taught and given practice using the mental processes (selective encoding, selective combination, selective comparison) alleged by the theory to be involved in figuring out meanings of new words from context.

2. Contextual-Cue training. Subjects were taught and given practice using the contextual cues upon which the mental processes operate (e.g., class membership, stative-descriptive).

3. Mediating-Variable training. Subjects were taught and given practice using the mediating variables that affect how well the processes can be applied to the cues (e.g., the location of a cue in the passage relative to the unknown word).

4. Vocabulary-Memorization control. Subjects were asked to memorize definitions of 75 extremely rare words (that otherwise did not appear in the experiment), and were tested on their memory for these words.

5. Context-Practice control. Subjects were given exactly the same practice that was given to subjects in the three training conditions, except that the practice occurred in the absence of training.

The mean pretest-posttest gain scores (out of 50 points possible on each test) were 7.2 for the process condition, 5.2 for the contextual-cue condition, 7.6 for the mediating-variables condition, 1.1 for the word-memorization control condition, and 2.6 for the context-practice condition. The results are clear: The training groups showed significantly greater gains than did the control groups. Two additional features of the means are worthy of note: First, as would be expected, the controls receiving relevant practice showed greater gain than did the controls receiving irrelevant memorization. The practice control condition is actually similar to many contextual training programs, which consist of little more than practice. Yet, to the extent that other programs involve any training at all, it is in contextual cues, which provide the least facilitation of all three training conditions.

In conclusion, theoretically-motivated instruction in learning words from context can make a significant and substantial difference in people's ability to learn word meanings, on their own, from context. In just 45 minutes of training, substantial gains in decontextualization ability were obtained. Of course, the durability of this training has yet to be shown.

Experiment 5. In all of the experiments reported up to this point, presentation of words was written, and subjects had as much time as they needed to read the passages and define the words. In everyday life, however, new words may be encountered in oral as well as in written presentations, and one may not always have as much time as one would desire to figure out the meanings of the new words. The present experiment addressed the question of what effects, if any, mode of presentation (oral, written) and rate of presentation (fast, slow) would have on quality of decontextualization. In particular, the experiment would allow a determination of whether subjects' decontextualization skills are in part a function of the medium and rate at which information is presented to the subjects.

Subjects were 62 Yale undergraduate and graduate students, equally divided between the sexes. Each subject received 15 passages of roughly 135 words per passage. Each passage had either two or four neologisms contained within it. Of the 15 passages, 5 dealt with literary or artistic topics, 5 with scientific topics, and 5 with history and current events. Subjects also received the Nelson-Denny Reading Test.

The main independent variables were type and rate of presentation. In the written-fast condition, subjects were allotted 45 seconds to read each passage. In

the written-slow condition, subjects were allotted 65 seconds per passage. In the oral-fast condition, passages were read aloud at a rate of 145 words per minute. In the oral-slow condition, passages were read aloud at a rate of 95 words per minute. Subjects were not allowed to look back in the written condition after they finished reading; similarly, replay of the tapes on which the oral presentations were made was not allowed in the oral condition.

Quality of definitions was rated on a 0 (low) to 2 (high) scale. Consider first the effect of mode of presentation. Mean quality ratings were 1.56 for the written conditions and 1.59 for the oral conditions--the difference between means was not significant. Consider next the effect of rate of presentation. Mean quality ratings were 1.67 for the slow-presentation condition and 1.48 for the fast-presentation condition--*this difference was significant.* The interaction between mode and rate was not significant. Hence, rate, but not mode of presentation, affected quality of decontextualization.

The correlations across items of each of the four conditions with each other were generally in the range from .6 to .8. Thus, the various conditions were similar, but not identical, in what they measured. Correlations with the Nelson-Denny Reading Comprehension score were .61 for the written-slow condition, .77 for the written-fast condition, .64 for the oral-slow condition, and .45 for the oral-fast condition. Thus, the condition most resembling that of a standardized reading test--the written-fast condition--showed the highest correlation with such a test. As would be expected, the written conditions combined correlated more highly with the Nelson-Denny scores (.63) than did the oral conditions combined (.45).

Experiment 6. In the previous experiment, the written and oral conditions did not differ in the quality of decontextualization they afforded, but this lack of difference may have stemmed from subjects in the written conditions not being allowed to look back at the passages after they finished reading them. Although such a manipulation increases the comparability of the written and oral conditions, it is perhaps less representative of most everyday reading than is a condition in which subjects are allowed to look back at the reading passage in order to figure out the meaning of a new word. A further experiment was thus conducted in order to determine the effects of looking back.

Sixty Yale students participated in an experiment that was parallel to the preceding one, except that there were two main manipulations: rate of presentation (fast versus slow) and lookback (allowed or not allowed). Presentation of passages was written only. As before, slow passage presentation resulted in better performance (mean quality score = 1.38) than did fast passage presentation (mean quality score = 1.34). The effect of lookback was clearcut: Subjects allowed to look back at passages in defining the new words had a mean *definitional quality* score of 1.53. Subjects not allowed to look back had a mean quality score of just 1.19. There was no interaction between rate and lookback conditions. Thus, lookback does facilitate decontextualization.

Correlations across item types between condition were in the range from .7 to .9. Thus, the various conditions were measuring similar skills. Correlations with the Nelson-Denny Reading Comprehension score were moderate, but the pattern was perplexing: .50 for slow look-back, .35 for fast look-back, .68 for slow no-look-back, and .48 for fast no-look-back. The correlations showed, if anything, the opposite to the pattern one might have predicted on the basis of surface similarities and dissimilarities of the experimental tasks to the Nelson-Denny test.

Theory of Decoding of Internal Context

Subjects use more than external context to figure out meanings of previously unknown words. *They use internal context as well.* By internal context, I refer to

the morphemes within a word constituted of multiple morphemes that combine to give the word its meaning. People attempting to figure out meanings of words will often use not only external context of the kinds discussed above, but internal context deriving from their prior knowledge of a new word's constituent morphemes. Together, the two kinds of context provide a potentially powerful set of clues for figuring out meanings of new words.

Research on the use of internal context has proceeded along somewhat different lines from research on the use of external context: The major theoretical issue seems to have been whether affixed words (such as predisposed) are stored in memory in unitary form (i.e., as predisposed), in a set of lexically decomposed forms (i.e., as pre-dis-pose-d), or in both of these kinds of forms, with the form that is accessed in a given instance depending upon the task and task context. From the point of view of a theory of verbal comprehension, research on internal context is less advanced than research on external context, in that the question of how internal context is used by comprehenders in vocabulary development has barely begun (but see Freyd & Baron, 1982), whereas the question of how external context is used in vocabulary development has received at least modest attention, as shown above. But recent research on how affixed words are stored in memory is relevant to the concerns of this chapter: If such words are stored only as single lexical entries, then one might expect comprehenders to have considerable difficulty in the use of internal context for inferring word meanings; if, on the other hand, affixed words are stored as sets of separate morphemes, either instead of or in addition to lexical entries for the complete words, then one might expect comprehenders to be able to use internal context fairly freely in inferring word meanings.

Evidence for the view that affixed words are stored as sets of separate lexical entries corresponding to their individual morphemes can be traced back at least to Taft and Forster (1975). These investigators performed three experiments employing a lexical-decision task. In this task, subjects are presented with a string of letters and have to decide as quickly and as accurately as possible whether the letter string does or does not constitute a real English-language word. Three major results of interest emerged. First, nonwords that were stems of prefixed words (e.g., juvenate, which is the stem for rejuvenate) took longer to recognize as nonwords than did nonwords that were not stems of prefixed words (e.g., pertoire, for which the re- in repertoire does not function as a prefix). This result suggested that the nonword stem was directly represented in the lexicon and that one or more extra steps were needed to identify the stem as a nonword. Second, words that could occur both as free morphemes--which can stand by themselves as words (e.g., vent as a word)--and as bound morphemes--which cannot stand by themselves as words (e.g., vent when serving as the stem of the word invent)--took longer to identify as words when the bound form was more frequent in the English language than was the free form. This result suggested that the existence of vent (or any other comparable letter string) as a bound morpheme and possibly as a salient and separate nonword lexical entry in memory may interfere with or in some other way lengthen the latency for recognition of vent (or any other comparable letter string) as a free morpheme that can stand as a word in its own right. Third, prefixed nonwords took longer to identify as nonwords when they contained a real stem (e.g., dejuvenate, in which juvenate is a real stem) than did control nonwords that did not contain a real stem (e.g., depertoire, in which pertoire is not a real, i.e., separable, stem). This result again suggested that the stem may have been stored separately and hence interfered with recognition of the total letter string as a nonword. Taft and Forster presented a model of word recognition that

incorporated the notion that morphemes are each represented as separate lexical entries. Further support for a notion of separate storage of individual morphemes can be inferred from a study of Murrell and Morton (1974), although the purpose of this study and the theoretical framework in which it was placed were different from those that are of concern here.

The Taft and Forster view of storage of affixed words by their individual morphemes was challenged by Manelis and Tharp (1977), who interpreted data from two of their own experiments as supporting the view that words are stored only as single units. These investigators also employed a lexical-decision task. Certain controls were introduced in stimulus selection that had not been implemented in the Taft and Forster (1975) studies. Two main results emerged. First, in a crucial comparison of latencies for words that were affixed with latencies for words that were not affixed, no significant difference emerged for time to recognize the two respective kinds of letter strings as words. Although this result supported the notion that affixed and nonaffixed words are stored in the same way, other results from this experiment strongly suggested that the two kinds of words are not identically processed, so that there was at least some ambiguity in the results. Second, a test condition designed to encourage subjects to use lexical decomposition failed to show the significant difference that would have supported the conclusion that decomposition was used. On the basis of these results, the investigators accepted a simple model in which affixed words are stored unitarily in the same form that nonaffixed words are stored.

Subsequent research (Stanners, Neiser, & Painton, 1979; Taft, 1979) has attempted to reconcile the two views presented above by suggesting that both forms of representation may be used and accessed at different points in information processing or in different tasks or task contexts. Of particular interest in this latter regard is a study by Rubin et al. (1979), which was interpreted by its authors as suggesting that decomposition is used only if special strategies are evoked as a result of a stimulus list that is heavily biased in favor of affixed words. At this point, the issue of just what differences, if any, exist between the representation and processing of affixed words remains unresolved.

The studies described above have all used a lexical-decision task, or some variant of it, to test the form in which affixed words are represented in memory. Obviously, any one paradigm is limited in the information it can provide regarding a given psychological issue. Holyoak, Glass, and Mah (1976) and Kintsch (1974) have done experiments related to those reviewed above that address the issue of lexical decomposition from a more semantic point of view. These investigators used semantic-decision and memory experiments to investigate lexical decomposition. Because the hypotheses they investigated, and hence the experimental paradigms, did not bear on exactly the issues discussed above (namely, the representation and processing of affixed words), the research will not be described in detail. It is worth noting, however, that the results of these experiments generally suggest that the representation of complex words is unitary, and that lexical decomposition is not routinely done. But some of the experiments suggest that lexical decomposition is possible under at least some circumstances.

The theoretical ideas discussed above have not been directly embodied in tests of verbal comprehension. Such tests do not separately measure people's knowledge of prefixes, stems, and suffixes, or people's ability to integrate these kinds of knowledge. But in standard vocabulary tests, in which words are presented for definition in the absence of any external context, the use of internal context provides the only viable means of figuring out meanings of words that are unknown or scarcely known. Introspective reports of vocabulary-test takers suggest that

they use their knowledge of prefixes, stems, and suffixes to figure out meanings of at least some words. Internal context, like external context, can on occasion impede attempts to infer word meanings. For example, meliorate and ameliorate have essentially the same meaning, despite the addition of the prefix a- in the latter form.

Psychologists and educators interested in vocabulary training have recognized the importance of internal context in vocabulary-skills training programs. Both Johnson and Pearson (1978) and O'Rourke (1974), for example, have incorporated training on intraword cues into their vocabulary development training programs. Indeed, the phonics approach to reading instruction can be viewed as preparatory to a program of training students on the use of internal contextual cues.

To summarize, evidence regarding the representation and processing of affixed words is mixed. My own reading of the evidence is similar to that of Miller and Johnson-Laird (1976), who have suggested that the subjective lexicon of each individual is organized in terms of the critical morphemes of derived words, even though each word has its own entry. Such an organization would allow people to use lexical decomposition in inferring the meanings of new words, at the same time that such use would require a distinct extra effort on the individual's part. In terms of our present interest in the use of internal context in inferring word meanings, the result of previous investigations leave open the question of whether or under what circumstances individuals actually do use their knowledge of word stems and affixes to figure out the meanings of affixed or other complex words.

Context cues. Because internal context is much more impoverished than is external context, the diversity of kinds of cues is much more restricted (see, e.g., Johnson & Pearson, 1978; O'Rourke, 1974). The four kinds of cues constituting our scheme (Sternberg, Powell, & Kaye, 1983) are

1. Prefix cues. Prefix cues generally facilitate decoding of a word's meaning. Occasionally, the prefix has a special meaning (e.g., pre- usually means "before") or what appears to be a prefix really is not (e.g., pre- in "predation"); in these cases, the perceived cue may be deceptive.

2. Stem cues. Stem cues are present in every word, in the sense that every word has a stem. Again, such cues may be deceptive if a given stem has multiple meanings and the wrong one is assigned.

3. Suffix cues. Suffix cues, too, generally facilitate decoding of a word's meaning; in unusual cases where the suffix takes on an atypical meaning, or in cases where what appears to be a suffix really isn't, the perceived cue may be deceptive.

4. Interactive cues. Interactive cues are formed when two or even three of the word parts described above convey information in combination that is not conveyed by a given cue considered in isolation from the rest of the word.

The usefulness of these kinds of cues in decoding meanings can be shown by an example. Suppose one's task is to infer the meaning of the word thermoluminescence (see Just & Carpenter, 1980). The word is probably unfamiliar to most people. But many people know that the prefix thermo- refers to heat, that the root luminesce is a verb meaning "to give off light," and that the suffix -ence is often used to form abstract nouns. Moreover, a reasonable interpretation of a possible relation between thermo- and luminesce would draw on one's knowledge that heat typically results in some degree of light. Note that this cue derives from an interaction between the prefix and stem: Neither element in itself would suggest that the light emitted from heat would be a relevant property for inferring word meaning. These cues might be combined to infer (correctly) that thermoluminescence refers to the property of light emission from heated objects.

Again, we make no claims that this simple (and unoriginal) parsing of internal contextual cues represents the only possible classification scheme, although we think it probably represents one, but not the only, plausible parsing. Collectively, these kinds of cues provide a basis for a person to exercise his or her competence in inferring word meanings.

Mediating variables. Again, there exists a set of variables that mediate the usefulness of cues. Our model includes five variables that affect cue usefulness. These variables are similar but not identical to those considered for external context:

1. Number of occurrences of the unknown word. In the case of internal contextual analysis, the context cues are the same on every presentation of a given unknown word. However, one's incentive to try to figure out the word's meaning is likely to be increased for a word that keeps reappearing relative to one's incentive to try to figure out the meaning of a word that appears just once or a very few times.

2. Importance of the unknown word to understanding the context in which it is embedded. Again, a word that is important for understanding the context in which it occurs is more likely to be worth the attention it needs for decontextualization. One can skip unimportant words, and often does. As before, importance can be subdivided into the importance of the unknown word to the sentence in which it is embedded and the importance of the word to the paragraph in which it is embedded.

3. Density of unknown words. If unknown words occur at high density, one may be overwhelmed at the magnitude of the task of figuring out the meanings of the words, and give up this task. Yet, it is possible that the greater the density of unfamiliar words in a passage, the more difficulty the reader will have in applying the external context cues, and hence the more important will be the internal context cues. A high density of unfamiliar words may encourage word-by-word processing and a greater focus on cues internal to unfamiliar words. This mediating variable interacts with the next one to be considered.

4. Density of decomposable unknown words. Because internal decontextualization may not be a regularly-used skill in many individuals' repertoires, individuals may need to be primed for its use. The presence of multiple decomposable unknown words can serve this function, helping the individual become aware that internal decontextualization is possible and feasible. In this case, the strategy is primed by repeated cues regarding its applicability.

5. Usefulness of previously known information in cue utilization. Again, one's knowledge of words, word cognates, and word parts will play an important part in internal decontextualization. The sparsity of information provided by such cues (in contrast to external cues) almost guarantees an important role for prior information.

Knowledge-acquisition components. The knowledge-acquisition components relevant to decontextualization of internal context are the same as those for decontextualization of external context, and hence will not be repeated here, other than by name: selective encoding, selective combination, and selective comparison.

Data testing the theory of internal decontextualization. We have sought to study use of internal context in two experiments.

Experiment 7. Our goal in our first study of internal decontextualization (Kaye & Sternberg, 1983) was to determine the extent to which secondary-school and college students could derive the correct definitions of very low-frequency words on the basis of their knowledge of frequently used prefixes and stems. We

sought to determine whether these students were attending to either of the words' constituents (prefix or stem) while attempting to define the words. We also examined relationships between students' metacognitive knowledge of such words and their actual performance in defining them. Given the present state of research in this area, we felt there is a need to know whether individuals use internal context before examining in detail how individuals use such context. Thus, our study tested a prerequisite for our theory to be applicable, rather than the theory itself, which we plan to test in subsequent research.

We tested 108 students, of whom 58 were in secondary school (approximately equally balanced among grades 8, 10, and 11) and of whom 50 were undergraduates at a state university. Each subject was exposed to 58 prefixed words that were selected each to contain 1 of 15 commonly used Latin prefixes and 1 of 15 commonly used Latin stems. Because there were 15 different prefixes and 15 different stems, there were a total of 30 different individual word parts. Each prefix and each stem appeared in from two to six different words. All words were of very low frequency and of 2-3 syllables in length. Each subject received half of the words in a multiple-choice word-definitions task and the other half of the words in a word-rating task. Words presented in each of the two tasks were counter-balanced across subjects.

In the word-definitions task, each word was paired with four possible definitions, one of which was correct and three of which were incorrect. One of the incorrect definitions retained the meaning of the prefix only, one retained the meaning of the stem only, and one retained the meaning of neither the prefix nor the stem. An example of such a problem is

EXSECT

(a) to cut out	(totally correct)
(b) to throw out	(prefix only correct)
(c) to cut against	(stem only correct)
(d) to throw against	(totally incorrect)

In the (metacognitive) word-rating task, each word was paired with four questions querying subjects' assessments of the state of their knowledge and its relation to each of the words with which they were presented: (a) How familiar is the word? (b) How easily can you define the word? (c) How similar is the word to another word you have seen or heard? (d) How similar is the word to another word you can define? Subjects responded by circling numbers arrayed on a 7-point scale, with more positive responses to the questions answered in terms of higher-numbered values on the scale.

All subjects were also asked to rate the 30 word parts for their meaningfulness (i.e., familiarity of meaning). This task, which occurred at the end of the experiment, also involved a 7-point rating scale.

A *hierarchical multiple regression procedure* was used to predict scores both on the learning-from-context cognitive task and on the metacognitive ratings of words and word parts. In such a procedure, sets of independent variables are entered into the regression in a fixed order and in successive steps. The variables entered at the first level of the hierarchy were always "dummy variables" for age, test form (i.e., which set of test words was received in which task), and age x test form. This level was used to control for the effects of those variables that might be expected to affect test performance, but that were not directly relevant to the question of how subjects answered the test items or made their ratings. The particular independent variables entered at the second level of the hierarchy were the theoretically relevant ones, and varied from one regression to the next. Consider, for example, the question of whether subjects are using prefixes in order

to figure out word meanings. If they are, then performance on a word with a given prefix, such as ex in the above example, should be better predicted by performance on other words sharing this prefix than by performance on words not sharing this prefix. If subjects are not using prefixes, then performance on words sharing the same prefix should be no better a predictor of performance on the given word than should performance on words not sharing the same prefix. The same logic applies for word stems. Thus, a typical independent variable at the second level of analysis would be performance on other words sharing the same prefix (for determining whether subjects use prefixes in figuring out word meanings) or sharing the same stem (for determining whether subjects use stems in figuring out word meanings). The variable entered at the third level of the hierarchical regression was an interaction term between the variables at the first and second levels of analysis.

The results suggested that college students, but not high school students, were able to use internal context to help infer word meanings. Values of R (the correlation between predicted and observed scores on each of the test items) were generally statistically significant for college students but not for high school students. However, both high school and college students had accurate metacognitive knowledge, that is, their metacognitive knowledge was predictive of their cognitive performance (and vice versa). Significant values of R for the various regressions ranged from .53 to .78 with a median of .63. The pattern of results suggested that the word stem was the central focus for determining what each of the various words meant, with the prefix modifying this stem meaning. Interestingly, knowledge of prefixes was better than knowledge of stems, at least for our word sample. This result may be attributable to the much larger number of *stems than of prefixes in the language.*

Experiment 8. A follow-up experiment, done in collaboration with Daniel Kaye, involved 80 college students. The students were divided into four groups, which varied in the tests of decontextualization skill that were presented. Every test contained decomposable words, decomposable pseudowords, nondecomposable words, and nondecomposable pseudowords. The decomposable items had meaningful prefixes and stems; the nondecomposable items did not. The words were genuine, although of low frequency; the pseudowords, of course, were not genuine. As in the previous experiment, each test item had associated with it four answer options, one of which was correct, and the others of which were incorrect. For the decomposable words, one distractor was correct in stem only, one in prefix only, and one in neither stem nor prefix. Although there were a total of 240 items, each subject received only 180 of these. In each of the four conditions, one of the four types of test items was omitted. Hence, subjects received only three of the four item types--decomposable words, decomposable pseudowords, nondecomposable words, and nondecomposable pseudowords.

We found that we could obtain significant prediction of success on words with a given prefix from other words containing that prefix, but a different stem (median $R = .48$), and that we could obtain significant prediction of success on words with a given stem from other words containing that stem, but a different prefix (median $R = .45$). Thus, in this experiment with college students, it appears that internal context was consistently used to figure out word meanings.

In conclusion, the data collected to date indicate the usefulness of the theory of verbal decontextualization for understanding something of how individuals acquire their vocabularies. The theory can explain, at some level, both differences in difficulty of learning individual words (stimulus variance) and differences in individuals' abilities to learn words (subject variance). Differences in word

difficulty are understood in terms of differences in cue availability, applicability of mediating variables, and interactions between different cues and mediating variables. Differences between subjects are understood in terms of their differential ability to use selective encoding, selective combination, and selective comparison upon the cues, and in terms of differences in susceptibility to the mediating variables.

Theory of Information Processing in Real-Time Verbal Comprehension

Consider now how verbal comprehension skills are executed in real time. I will first describe two general alternative approaches to this issue, and then consider in more detail our own approach.

Alternative Approaches to Understanding Real-Time Verbal Comprehension

Approaches emphasizing current functioning seem divisible into two subapproaches--those that are essentially molar, dealing with information processing at the level of the word, and those that are essentially molecular, dealing with information processing at the level of word attributes. I shall consider each subapproach in turn.

A molar subapproach. The molar subapproach examines comprehension and understanding of individual words or groupings of words. A proponent of this approach, Marshalek (1981), administered a faceted vocabulary test along with a battery of standard reasoning and other tests. The facets of the vocabulary test were word abstractness (concrete, medium, abstract), word frequency (low, medium, high), item type (vague recognition--easy distractors in a multiple-choice recognition task, accurate recognition--difficult distractors in a multiple-choice recognition task, definition--subjects have to provide word definition rather than being given multiple-choice), and blocks (two parallel blocks of words). Marshalek found that vocabulary item difficulty increased with word abstractness, word infrequency, item formats requiring more precise discrimination of word meaning, and with task requirement (such that word definition was harder than word recognition). He also found that partial concepts are prevalent in young adults and that word acquisition is a gradual process. Vocabulary level seemed to be related to reasoning performance at the lower but not the higher end of the vocabulary difficulty distribution. These results led Marshalek to conclude that a certain level of reasoning ability may be prerequisite for extraction of word meaning (see also Anderson & Freebody, 1979). Above this level, the importance of reasoning begins rapidly to decrease.

Marshalek's approach to understanding verbal comprehension is of particular interest because it breaks down global task performance into more specific facets. It is possible, in his research, to score each subject for the various facets of performance as well as for the overall level of performance. I believe this to be an important step toward understanding current verbal functioning. One concern I have, though, is with whether the experimenter-defined facets correspond to important psychological (subject-defined) aspects of performance. Although these facets may differentiate more and less difficult items, and better and poorer performers, it is not clear that they do so in a way that bears any resemblance to the psychology of verbal comprehension. In other words, it is not clear how understanding these facets of performance gives us what could in any sense be construed as a causal-explanatory account of verbal comprehension and individual differences in it. The causal inferences that can be made are, at best, highly indirect.

A molecular subapproach. The molecular subapproach is the kind that we have taken in our work on the real-time representation and processing of information during verbal comprehension. The idea is to understand verbal

comprehension in terms of how attributes of words are encoded and compared, and also to understand decision-making in real-time reading through the specific decisions that are made about allocating time. For example, one would seek to understand performance on a synonyms test in terms of actual comparisons between the attributes of a given target word and the attributes of the potential synonyms given in a multiple-choice list. At minimum, one would have to know what kinds of attributes are stored, how these attributes are stored, how these attributes are accessed during verbal comprehension performance, and how these attributes are compared between the target and the options. Our theory of these phenomena (McNamara & Sternberg, 1983; Sternberg & McNamara, 1985), and some data testing the theory, are presented next.

Performance Components

In work investigating the performance components of real-time information processing, Timothy McNamara and I have sought to understand the mental representations and processes people use in understanding and comparing word meanings.

Alternative models of word representation. Several alternative models have been proposed for how word meaning is represented mentally. I consider below some of the major models that have been proposed.

Defining attribute (nonadditive) models. Traditional models of word meaning make use of necessary and sufficient--i.e., defining--attributes of words (Frege, 1952; Russell, 1956). The idea is that the meaning of a word is decomposed into a set of attributes such that the possession of these attributes is necessary and sufficient for a word to refer to a given object or concept. For example, a bachelor might be represented in terms of the attributes: unmarried, male, and adult. Being an unmarried male adult is then viewed as necessary and sufficient for being labeled as a bachelor. (Some might add never-before-married as an additional required attribute.) Traditional models can be viewed as "nonadditive," in the sense that either a given word has the attributes necessary and sufficient to refer to a given object or concept, or it does not; there are no gradations built into this model of representation.

Characteristic attribute (additive) models. A second class of models, and one that has been more in favor in recent times, might be referred to as "characteristic attribute" models. In these models, word meaning is conceptualized in terms of attributes that tend to be characteristic of a given object or concept, but neither necessary nor sufficient for reference to that concept. A well-known example of the usefulness of this kind of model stems from Wittgenstein's (1953) analysis of the concept of game. It is extremely difficult to speak of necessary attributes of a game. Similarly, it is difficult to speak of any attributes that guarantee something's being a game: Hence, it is difficult to find any sufficient attributes of a game. Yet, games bear a "family resemblance" to each other. In today's parlance, various games cluster around a "prototype" for the concept of a game (Rosch, 1978). Games are either closer to or further from this prototype depending upon the number of characteristic attributes of a game they have. A game such as chess might be viewed as quite close to the hypothetical prototype, whereas a game such as solitaire might be viewed as further away from this hypothetical prototype.

The class of additive models can be divided into at least three submodels according to how the attributes are used to refer to a concept: (a) The reference of a word might be determined by the number of attributes possessed by an object that match attributes in the word's definition (Hampton, 1979; Wittgenstein, 1953). If the number of matching attributes exceeds some criterion, then the object is

identified as an example of the word; otherwise, the object is not so identified. (b) The referent of a word might be determined by a weighted sum of attributes. This model is like the first one, except that some attributes are viewed as more critical than are others, and hence are weighted more heavily (Hampton, 1979). For purposes of our analyses, the first model will be viewed as a special case of the second (the weighted case), and will not be treated as qualitatively distinct. (c) The referent of a word might be determined by a weighted average of attributes, in which case the sum of the attributes is divided by the number of attributes. The second and third models are distinguished by whether or not a given sum of weights counts equally without respect to the number of weights entered into the sum. To our knowledge, the difference between summing and averaging models has not been addressed in the literature on word meaning and reference, although it has certainly been considered in other contexts, such as information integration in people's formation of impressions about each other (e.g., Anderson, 1979).

Mixture models. A third class of models specifies words as being decomposable into both defining and characteristic attributes. An example of such a model would be that of Smith, Shoben, and Rips (1974), who proposed that words can be viewed as comprising both defining and characteristic attributes. Consider, for example, the concept of a mammal. Being warm-blooded would be a defining attribute of a mammal, whereas being a land-animal would be a characteristic attribute, in that most, but not all, mammals are land animals.

In the mixture model (or at least the proposed variant of it), not all words need be composed of both defining and characteristic attributes (Clark & Clark, 1977; Schwartz, 1977). For example, one might view some words, such as game, as comprising only characteristic attributes. Intuitively, it seems much easier to find defining attributes for some kinds of concepts than for others, and this class of models capitalizes upon this intuition. It seems less likely that any words comprise only defining attributes. At least, we are unable to think of any words that do not have at least some characteristic attributes that are neither necessary nor sufficient for referring to a concept.

Tests of alternative models of representation. We conducted four initial experiments to test the alternative models of word-meaning representation (McNamara & Sternberg, 1983). Our concern in these experiments was with how word meaning is represented psychologically. The psychological issues of interest to us are not, of course, necessarily the same as those issues concerning philosophers of meaning and linguists.

The first experiment was intended to (a) determine whether people identify necessary and/or sufficient attributes of concepts and objects and (b) collect rating data needed for a second experiment that tested the various models of representation. Ten Yale students participated in the study. The study involved three kinds of nouns: (a) natural-kind terms (e.g., eagle, banana, potato), (b) defined-kind terms (e.g., scientist, wisdom), and (c) proper names (e.g., Queen Elizabeth II, Aristotle, Paul Newman). Proper names were included because they have been heavily used in the philosophical literature, often serving as the basis for generalization to all nouns. The main independent variables in the experiment were the type of term about which a rating was to be made (natural kind, defined kind, proper name) and the type of rating to be made (necessary attributes, sufficient attributes, importance of attributes--see below). The main dependent variable was the value of the assigned ratings. Subjects were first asked to list as many properties as they could think of for the various objects of the three kinds noted above. Then they were asked to provide three kinds of ratings (with the order of the kinds of ratings counterbalanced across subjects). The first kind of

rating was one of necessity: Subjects were asked to check off those attributes, if any, that they believed to be necessary attributes for each given word. The second kind of rating was one of sufficiency: Subjects were asked to check off those attributes, if any, that were sufficient attributes for each given word. They were also asked to indicate minimally sufficient subsets of attributes (such that the subset in combination was sufficient to define a word). In both of these kinds of ratings, it was emphasized that there might well be no necessary or sufficient properties (or subsets of properties) at all. The third kind of rating was one of importance: Subjects were asked to rate how important each attribute was to defining each of the given words. These ratings were used to determine how characteristic each attribute is of the concept it helps describe.

The major results were these:

First, all subjects found at least one necessary attribute for each of the eight natural kind and proper name terms. All but one subject found at least one necessary attribute for each of the defined kinds. One could therefore conclude that individuals conceive of words of these three kinds as having at least some necessary attributes. Examples of some of these attributes are, for a diamond, that it scratches glass, is the hardest substance known, and is made of carbon; and for Albert Einstein, that he is dead, was a scientist, was male, and that he invented the equation $E = MC^2$.

Second, all subjects found at least one sufficient attribute or subset of attributes for all natural kind terms. Almost all subjects found at least one sufficient attribute or subset of attributes for defined kinds and proper names. One could therefore conclude that most individuals conceive of most words as having at least some sufficient attributes or subsets of attributes. Examples are, for an eagle, that it is a bird that appears on quarters, and for lamp, that it is a light source that has a shade.

Third, roughly half of the natural kind and defined kind terms were conceived as having attributes that were both necessary and sufficient. More than three-fourths of the proper names were conceived as having such attributes. Examples are, for sandals, that they are shoes that are held on with straps and that do not cover the whole foot, and for a diamond, that it is the hardest substance known.

Fourth, internal-consistency analyses revealed that subjects agreed to a great extent as to what attributes were important, necessary, sufficient, and necessary and sufficient (with internal-consistency reliabilities generally in the mid-.80s; for necessity ratings and sufficiency ratings, reliabilities were generally a bit lower, usually in the mid-.70s).

The second experiment was intended to (a) determine the extent to which people use defining (necessary and sufficient) and characteristic (neither necessary nor sufficient) attributes when deciding whether or not an object is an exemplar of a word, (b) to test four simple models and three mixture models of word meaning, and (c) to determine how generalizable the results were across word domains. Nine of the ten subjects from the first experiment participated in this experiment. A within-subjects design was used in order to control for possible individual differences in the representation of meaning of specific words. The subjects received booklets with a given word at the top of the page, followed by a list of attributes. The subject's task was to give a confidence rating that the attributes actually described an exemplar of the word at the top of the page. Attribute descriptions were compiled for each subject in order to provide discrimination among alternative models of word representation. The main independent variables were ratings of necessity, sufficiency, necessity and sufficiency, and importance, as taken from Experiment 1. The main dependent variable was the confidence

rating that the description described an exemplar of the target word. Subjects rated their confidence that a given word was, in fact, exemplified by the description appearing below it. For example, one might see the word TIGER at the top of the page, followed by four attributes: "member of the cat family," "four-legged," "carnivorous," and "an animal." One would rate on a 1-8 scale how likely that list of attributes was to describe a particular tiger.

The alternative representational models tested were a model positing (a) use only of defining (necessary and sufficient) attributes; (b) use of an unweighted sum of attributes; (c) use of a weighted sum of attributes; (d) use of a weighted mean of attributes; (e) use of defining attributes as well as a weighted sum of all attributes; and (f) use of defining attributes as well as a weighted mean of all attributes. Models were fit by linear regression with individual data sets concatenated; that is, there was no averaging across either subjects or items, and thus there was just one observation per data point for a total of 863 data points. Proportions of variance accounted for by each of the six respective models in the confidence-rating data were .36 for (a), .01 for (b), .02 for (c), .11 for (d), .45 for (e), and .38 for (f), concatenated over word types. Data for individual subjects reflected the pattern for the group. It was concluded that in making decisions about whether sets of attributes represent exemplars of specific words, individuals appear to use both defining and characteristic attributes via the weighted sum model.

The third experiment was parallel to the first, in that it replicated this experiment and also provided needed ratings data for the subsequent experiment. Because the results were almost identical to those of the first experiment, they will not be presented separately here.

The fourth experiment was designed to verify the results of the second experiment using converging operations. In particular, response latency and response choice were used as dependent variables, and the subjects' task was to choose which of two attribute lists better described a referent of a given word. For example, subjects might see "SOFA," followed by two lists of attributes: (1) "used for sitting, found in living rooms, slept on, furniture," and (2) "slept on, rectangular in shape, found in bedrooms." The 32 subjects would have to decide whether (1) or (2) was a better exemplar of sofa. Models were fit to group-average data. In this experiment, as in the previous two, natural kinds, defined kinds, and proper names appeared in equal numbers as stimulus terms.

The results again supported the mixture model combining defining attributes with summed characteristic attributes. For response choices, fits of five of the models described earlier were .48 for (a), .57 for (c), .46 for (d), .65 for (e), and .57 for (f). Model (b), the unweighted variant of Model (c), was not separately tested.

The data for the four experiments taken as a whole seemed to support the mixture model in which defining attributes and characteristic attributes are considered, with the former attributes considered both nonadditively and as a weighted sum combined with the latter attributes. This model was then taken as the representational model on the basis of which to test a process model.

Model of information processing. We have proposed a model that assumes that, in Experiment 4, (a) subjects tested both answer options in order to make sure that they picked the better of the two options and (b) subjects compared answer options on the basis of both defining attributes (when present) and weighted sums of attributes. A flow chart for the model can be found in Figure 3.

(INSERT FIGURE 3 ABOUT HERE)

Quantification of the model. Quantification of the processing model, which serves as a basis for testing the processing model, will be explained by referring to the following stimulus item: "TENT" followed by (1) "Made of canvas, supported by poles, portable, waterproof" and (2) "A shelter, used for camping, made of canvas." The six parameters of the model and the variables used to estimate them were as follows:

(1) Reading time was estimated by the total number of words in the two descriptions, excluding the target word. In the example, the total number of words is 16. The value of this variable ranged from 4 to 34 across items (mean = 14.0).

(2) Processing time for negations was estimated by the number of negated attributes in the descriptions, which is 0 for the example. The value of this variable ranged from 0 to 2 across items (mean = 0.6).

3. Time for comparison of attributes in the descriptions to attributes of the target word was estimated by the total number of attributes in the two descriptions. According to the model, each attribute in each description is compared to the attributes of the encoded target word. The weights of matching and mismatching attributes are added to a weighted-sum counter for the description currently being processed (there is a weighted-sum counter for each description). Mismatching attributes are also checked for necessity, and if they are necessary, this information is recorded in a defining-attributes counter for the description currently being processed (there is a defining-attributes counter for each description). When all atttributes in a description have been compared to the attributes of the target word, the description is checked for sufficiency. If the description is sufficient, this fact is recorded in the defining-attributes counter for that description. In the example, the comparison variable would take the value 7, the number of attributes in the two descriptions. The comparison variable ranged from 3 to 8 across items (mean = 5.7).

4. Comparison of options on the basis of defining attributes was estimated by the absolute difference between the number of subjects for whom the second description was sufficient and the number of subjects for whom the first description was sufficient, or by the absolute difference between the number of subjects for whom a negated attribute in the first description was necessary and the number of subjects for whom a negated attribute in the second description was necessary. According to the model, comparison time decreases as the difference between the defining-attributes counter for the first description and the defining-attributes counter for the second description increases; i.e., subjects are faster the more dissimilar the options are. We needed to use a continuous variable to estimate a dichotomous construct (the necessity or sufficiency of a set of attributes) because we were modeling group-average data and a given description was not equally good or bad for all subjects. The difference between the two descriptions capitalized on this inherent variability in our stimuli. This comparison variable was linearly scaled so that small values on the variable corresponded to large differences between the two descriptions, and hence to fast comparison times. In the example, the first description was sufficient for none of the subjects and the second was sufficient for 11 subjects. Thus, the comparison variable was a linear function of the number 11 (precisely 26-11, or 15).

5. Comparison of options on the basis of weighted sums of attributes was estimated by the absolute difference in summed weights between the two descriptions. It was assumed that comparison time decreases as the difference between the weighted-sum counter for the first description and the weighted-sum counter for the second description increases. This variable, like (4), was linearly scaled so that small values on the variable corresponded to large differences between the two descriptions. In the example above, the first description had a weighted sum of 11.32 and the second description had a weighted sum of 11.44. Hence, the comparison variable was a linear function of 0.12 (precisely 17.76-0.12, or 17.64).

6. Justification was relevant when the difference in summed weights and the difference in defining attributes predicted opposite choices. In such cases, the choice of an answer option on the basis of definingness alone had to be justified. The stimulus items were constructed so that for 60 of the 156 pairs of descriptions, there were no differences in definingness between the descriptions. For these items, the justification variable always took the value 0, since there could be no discrepancy between choices. The difference in weighted sums and the difference in definingness predicted opposite choices for 10 of the remaining 96 items. For these items, the justification variable took the value 1. In the example, both the difference in definingness and the difference in weighted sums predict that the second options should be chosen. Thus, the value of the justification variable is 0.

Tests of model of information processing. The model described above was tested in terms of its ability to account for mean response latencies on the 156 items. Responses were included in the mean latencies even if they were errors according to the model. Fits changed trivially when errors were excluded. Fit of the model (R^2) was .79, with an RMSD of .66 second. Standardized parameter estimates were .42 for reading time, .37 for processing negations, .33 for comparison to target word, .23 for definingness comparison, .33 for weighted-sum comparison, and .07 for justification. All estimates were statistically significant. The proposed model thus provided a good account of the processing of attribute information, accounting for nearly 80% of the total variance in response latencies (and with only six independent variables on 156 data points). The standardized regression coefficients for the model seemed generally reasonable. They indicated that weighted sums of attributes were somewhat more important than defining attributes in deciding which option was the better exemplar of the target word.

Correlations with Ability Tests

Correlations were computed between overall mean latencies on the decision task and scores from the Nelson-Denny Reading Test and the Differential Aptitudes Test (DAT). In particular, we used vocabulary, reading comprehension, and reading rate scores from the former test, and the verbal reasoning score from the latter test. The only significant correlation involving latency was that between overall mean latency and reading rate (-.37). However, the multiple correlation between mean latency, on the one hand, and both reading rate and comprehension, when taken together, was a significant .47. (The correlation between comprehension and reading rate was .07, and the correlation between comprehension and mean latency was .27; neither correlation was significant.) Both reading rate and comprehension made statistically significant contributions to the multiple correlation, with respective weights of .38 and .30. Thus, reading rate and comprehension, when considered together, were moderately strongly related to latency in our task.

To conclude, the results from the four experiments taken together provide a reasonably coherent picture of both the representation of meaning and the processes used to make reference to a concept. In particular, individuals seem to use an additively-based mixture model in their representation of word information, and to be able to combine the represented information in a way that enables them to choose synonyms. Obviously, our work on real-time processing is incomplete. It has yet to be extended beyond the level of individual words, and is in need of further interface with the theory of learning from context. Nevertheless, the two aspects of the theory of verbal comprehension, taken in combination, seem to provide a relatively comprehensive view of how crystallized intelligence develops and functions.

Metacomponents

Virtually everyone is confronted with much more material to read than they could possibly handle in the time allotted for reading. College freshmen, for example, are often bewildered by a reading load that would seemingly take all of their time if they were to attempt to read with care all of the material they are assigned. Professionals in psychology and other fields often find it impossible to keep up with the latest developments in their field simply because there are so many of them to read about, but not nearly enough time to do the required reading. Clearly, individuals have to allocate their reading time and depth of reading in a way that reflects the realities of their usually overburdened situation.

A number of investigators have recognized the importance of metacomponents, or executive processes, in reading. For example, Brown (1978, 1980) has proposed that metacognitive skills operate in conjunction with an "automatic" mode of reading. Taking a developmental approach, Brown has reported that young children are deficient at, among other things, (a) predicting the difficulty of a task and recognizing when task difficulty has changed (see, e.g., Salatas & Flavell, 1976); (b) comprehension monitoring--being aware of whether one does or does not understand (see, e.g., Markman, 1977, 1979); (c) study-time apportionment--studying in anticipating of a future test--which includes determining what is important to remember and what is not, choosing a strategy to maximize learning, determining one's success with the chosen strategy, and determining whether a different strategy should be tried (see, e.g., Masur, McIntyre, & Flavell, 1973); and (d) predicting test performance--knowing when a task has been mastered.

A strength of this approach is that executive processes provide a means for accounting for many of the complexities of skilled reading. To date, most of the evidence for the importance of executive control in cognitive performance resides largely in comparisons of performance across groups differing substantially in developmental level. Deficiencies in executive skills have been shown to be responsible for at least some limitations in cognitive performance that are characteristic of young children. The importance of efficient executive functioning to adult skilled performance remains largely an open question (Brown, 1980).

Richard Wagner and I have conducted two experiments in order to investigate executive processes in reading of expository texts by adults (Wagner & Sternberg, in press). This research derives from the issues raised above.

Isolating a time-allocation metacomponent. In a first experiment, subjects were 40 Yale undergraduates. Each subject was presented with 44 untitled passages of about 150 words apiece. One-fourth of the passages were from novels, one-fourth from newspapers, one-fourth from humanities textbooks, and one-fourth from science (natural and social) textbooks. Although there were eight different questions per passage, a given subject saw just two of these. These two questions addressed either the gist of the passage (i.e., general points), the main idea of the passage, specific details in the passage, or analysis and application of points in the passage (i.e., inferences and evaluations from the text). Which subjects received which questions for which passages was counterbalanced across subjects. Subjects also received the Nelson-Denny Reading Test and the Differential Aptitude Test Verbal Reasoning (verbal analogies) subtest.

Subjects received 11 trials of 150 seconds each. Each trial involved 4 reading passages, for a total of 44 passages. Subjects were informed for each passage of whether they would be tested for gist, main ideas, details, or analysis and application. Subjects were free to allocate their total time across passages as they wished. Passages were selected by subjects pressing an appropriately designated

key on a computer console. Thus, order of presentation of passages within trial, and duration of viewing, were under subject control. Note, then, that subjects were basically free to allocate their time to the four types of questions as they wished.

Mean latencies for passages read for each of the different purposes (question types) were 38.0 for gist, 37.6 for main idea, 39.8 for details, and 40.4 for analysis and application. The times differed significantly from each other, with the times for gist and main idea comprehension significantly shorter than the times for detail comprehension and analysis and application comprehension. Thus, subjects did allocate time systematically so as to spend more time reading passages for which they would receive more demanding questions. Patterns of accuracy in responding to the question types also were systematic: Mean numbers of questions answered correctly (out of 16) for each type of question were 13.3 for gist, 12.6 for main idea, 10.5 for details, and 8.2 for analysis and application. These means, too, differed significantly from one another. The means for gist and main idea comprehension were significantly higher than those for details, which in turn was significantly higher than for analysis and application.

Overall number of questions correctly answered by each subject for all types of questions was significantly correlated with vocabulary (.57), comprehension (.48), and DAT verbal reasoning (.78). Most of the subscores were also significantly correlated with Nelson-Denny and DAT scores, and all correlations were in the predicted (positive) direction. Thus, our reading questions did seem to measure skills related to those measured by standard tests of reading comprehension.

The most important question, from our point of view, was that of whether time allocation was systematically related to task performance. A time-allocation score was computed for each subject by subtracting the amount of time spent on reading passages for gist and main idea from the amount of time spent on reading passages for details and analysis and application. Presumably, a higher difference score would reflect greater sensitivity in time allocation: The higher the score, the relatively greater the amount of time spent on reading for the more difficult questions and the relatively lesser the amount of time spent on reading for the less difficult questions. Time allocation score correlated .30 with total number of passage comprehension questions answered correctly. But one might well ask whether this correlation merely reflects some skill already measured by standard reading comprehension tests, which seem to measure primarily performance components rather than metacomponents in reading. We therefore predicted accuracy in answering questions on the reading task from DAT Verbal Reasoning score, Nelson-Denny total score (reading comprehension + vocabulary), Nelson-Denny reading rate, and time allocation. The question addressed was whether time allocation would make a significant contribution to the regression after the other, standard test variables were added to the equation. In fact, it did. The semi-partial regression weight for the time allocation parameter was .30, which was statistically significant. The overall multiple correlation was .85. Thus, our metacomponential measure of reading time allocation makes a significant contribution in predicting task performance over and above that made by standardized test scores (including vocabulary, comprehension, verbal reasoning, and reading rate). Again, metacomponential processing seems to be important in real-time verbal comprehension.

Strategies in using adjunct information. In a second experiment, 90 Yale undergraduates were divided into three groups.

In a control group, subjects received 8 passages and 44 questions taken from the reading comprehension sections of two editions of the Graduate Record

Examination of the Educational Testing Service. The passages were of two lengths. Four of the passages were approximately 175 words in length. There were three questions on each of these passages. The other four passages were approximately 500 words in length. There were eight questions on each of these passages.

In a *difficulty-information group*, subjects received the identical passages and questions to those used in the control group, but with two kinds of difficulty information added. General difficulty information informed subjects of the average difficulty of the set of questions associated with each passage. This information was conveyed through a table containing the average difficulty of the questions associated with each passage and the number of questions per passage (either three or eight questions per passage). Specific difficulty information informed subjects of the difficulty level of each question. This information was conveyed by labeling each question on a scale of relative difficulty. Both general and specific difficulty information was presented through use of the phrases "very difficult," "moderately difficult," "moderately easy," and "very easy." Difficulty level was determined by the proportions of the examinees who passed the question when it was administered nationwide as part of the Graduate Record Examinations.

In an importance information group, subjects received the identical passages and questions to those received in the other two groups (but without any difficulty information). However, the most important sentences in the passages (as determined by the judgment of the experimenters) were highlighted with a yellow marker pen. Approximately 45% of the text was highlighted.

Subjects in the difficulty information condition were instructed to use the difficulty information so as to maximize their performance. They were not told how to do so, however. Examples of questions labeled as to their difficulty were provided. Subjects in the importance information condition were instructed to use the importance information so as to maximize their performance. Again, they were not told how to do so. An example of a highlighted text was provided.

In all conditions, subjects reported the time (from an easily visible clock) that they started work on each reading passage. At the conclusion of the task, subjects in all groups provided written descriptions of their task strategies. Subjects in the two experimental conditions additionally described whether they made use of the adjunct difficulty or importance information, and if so, how.

Reference ability measures of reading and reasoning abilities were the same as those employed in the previously described experiment.

Task performance in terms of total number of questions (out of 44 total) answered correctly was 25.7, 23.6, and 21.6 for the control, difficulty information, and importance information conditions. These means did not differ significantly. Written reports of task strategies provided by subjects were scored for presence of explicit mention of revising strategy during task performance. Twenty-eight percent of subject reported strategy revision during task performance. This percentage remained essentially constant across conditions: 30%, 30%, and 23% for the control, difficulty information, and importance information conditions, respectively. Two reasons were given for strategy revision: a strategy chosen before beginning the task was not working out, or a strategy was changed when time began to run out so that the subject would have a chance of answering the remaining questions. Subjects who reported strategy revision performed significantly better on the task than subjects who did not, achieving a mean total score of 26.1, compared to 22.7 for subjects who did not report strategy revision. Subjects who reported strategy revision also attained significantly higher scores on the DAT Verbal Reasoning Test than did subjects who did not (45.7 versus 44.1); scores on the Nelson-Denny Reading Test did not differ significantly between

groups, however.

Recall that subjects marked down the time when they began work on each passage, producing a record of the order in which passages were read. This record was used to score the presence or absence of the strategy of reading passages in their order of difficulty. No attempt was made to distinguish between subjects who followed this strategy exclusively and subjects who followed this strategy for only part of the reading task. Fifty-three percent of subjects in the difficulty condition used this strategy. Subjects using this strategy obtained higher average task scores than subjects who did not, 26.3 versus 20.6. Subjects using this strategy also obtained marginally significantly higher scores on the Nelson-Denny Reading Test (136.3 versus 123.9), and had a higher reading rate as measured by the Nelson-Denny (348.8 versus 286.8). Performance on the DAT Verbal Reasoning Test did not differ between the two types of subjects (43.9 versus 44.1). More able subjects, *then, used general difficulty information in planning their order of passage reading* to correspond with the order of passage difficulty. It was possible to determine the validity of subjects' written reports of task strategies by comparing actual strategy as determined from the record of passage order with written reports of strategy. All subjects who used the strategy of reading passages in order of their difficulty reported doing so; conversely, no subjects who did not use this strategy reported doing so.

Subjects' written reports of task strategy were scored for (a) the presence of a strategy of using the specific difficulty information and for (b) indications that the specific information was distracting. Twenty-seven percent of subjects reported using the specific difficulty information. Subjects who reported using the specific difficulty information described a strategy of matching how much effort they spent searching for and evaluating possible answers to the difficulty level of the questions. Subjects who reported using the specific difficulty information actually performed more <u>poorly</u> on the task than did subjects who did not report use of specific difficulty information (18.8 versus 25.4). These subjects also obtained lower Nelson-Denny Reading Test scores (111.0 versus 136.5), but performance on the DAT Verbal Reasoning Test did not differ significantly across groups (41.6 versus 44.8). Twenty percent of subjects reported that the specific difficulty information was distracting. One commonly given reason for the general unhelpfulness of the specific difficulty information was that the difficulty rating of a particular question did not coincide with a subject's personally perceived difficulty. Subjects also reported that they disliked being told how difficult a question was; in some cases, knowing that a question was very difficult made them anxious. Subjects who reported the specific difficulty information as distracting performed better on the task than did subjects who did not (31.2 versus 21.8). No reliable differences were found on the reference ability tests, however, for this comparison. Overall, then, more able subjects were (a) more likely to use <u>general</u> difficulty information for planning order of passage selection, were (b) less likely to use <u>specific</u> difficulty information, and were (c) more likely to find the <u>specific</u> difficulty information distracting.

Three strategies for using importance information were identified. Twenty-seven percent of subjects reported using a strategy of reading highlighted sections exclusively. Task performance for subjects using this strategy was comparable to that of subjects who did not use the strategy (21.8 versus 21.6), as was performance on the Nelson-Denny Reading Test (125.0 versus 130.1). Subjects who used this strategy did perform better on the DAT Verbal Reasoning Test, however (46.9 versus 44.3). A second strategy, related to the previous one, was reading the highlighted sections more carefully than the non-highlighted sections, but not

exclusively. Forty-three percent of subjects reported using this strategy, but performance on the task and reference ability measures was comparable for subjects who reported using this strategy and those who *did not*. A *final* identifiable strategy was one of searching for unknown answers in the highlighted portions. This was an understandable strategy because a majority of the answers were to be found in the highlighted sections and subjects were informed of this fact. Thirty-three percent of subjects reported using this strategy. Task performance was comparable for subjects who reported using this strategy versus those who did not (23.8 versus 20.6). Subjects who reported using this strategy performed better on the Nelson-Denny Reading Test, however (139.8 versus 123.7). Their performance on the DAT Verbal Reasoning Test was comparable to that of subjects who did not use this strategy (46.1 versus 44.4).

To conclude, the proposed theory of real-time verbal comprehension appears to give a good account of processing both at the word level and for time allocation at the passage level. The theory is by no means a complete theory of real-time processing, but may at least provide a step in that direction. In combination, the theory of how verbal comprehension develops and the theory of how verbal comprehension functions in real time seem to provide a reasonable start toward understanding the psychology of verbal comprehension.

<u>Learning and Reasoning with Novel Concepts</u>

All of the studies reported above have dealt with acquisition and processing of fairly conventional kinds of words and concepts. However, according to the triarchic theory of human intelligence, which motivates my research on human intelligence, an important part of verbal intelligence is the ability to acquire and process novel, or nonentrenched concepts. In research supported by my previous ONR contract, I showed that it is possible to isolate people's ability to reason with nonentrenched concepts (see Sternberg, 1982, 1985). But this research did not provide a clear idea of just what nonentrenchment is, and particularly, of whether it is a property of verbal <u>concepts</u>, of verbal <u>tokens</u> (i.e., words), or some combination. A further set of studies was undertaken by Tetewsky and Sternberg (1985) in order to clarify the nature of nonentrenchment.

A reasoning task was designed in which the naturalness of a concept and the type of name used to define that concept could be independently varied. A conceptual system was found in which the content could be expressed in four different forms, so the two levels of concepts (natural or unnatural) could be crossed with two levels of names (familiar or novel). The underlying assumption for this design was that these two variables might be important in distinguishing between entrenched and nonentrenched concepts.

In the first experiment, subjects were required to solve reasoning problems in which they had to select among alternative projections about occurrences in the environment that relate to seasonal changes. It is quite natural for the leaves to change color in accordance with the seasons (at least in New England!). However, it is not at all natural for us to think that rocks will change color according to a seasonal pattern. Analogously, seasons can be identified by the names <u>summer</u>, <u>fall</u>, <u>winter</u>, and <u>spring</u>, or they can be given novel names, such as <u>soob</u>, <u>trit</u>, <u>blen</u>, and <u>mave</u>. By using these two sets of concepts and names, the following four situations were constructed: (a) familiar season names describing states of the leaves; (b) novel season names describing states of the leaves; (c) familiar season names describing states of the rocks; and (d) unfamiliar season names describing states of the rocks. In the second experiment, subjects were required to make projections about events in the environment that relate to periods of the day. In this context, it is natural to identify a period of the day by noting the position of

the sun relative to the horizon and it is quite unnatural to expect that minerals will change shape as the day progresses. Also, the periods of the day can be identified by the names daytime, dusk, nighttime, and dawn, or they can be given novel names such as trolar, bren, stobe, and kovit. By using these two sets of concepts and names, a set of four situations was constructed that was structurally equivalent to that describing seasons in the first experiment.

Subjects in each of these conditions were given descriptions of the beginning and end of a season (or the beginning and end of a period of the day) and were required to make inferences regarding the events that occurred. The problems were presented individually as "selection task items." The ease with which subjects made these judgments was measured by both latency and error indices. A model of information processing was also tested for the latency data obtained in each of the four tasks. In addition to providing an empirical analysis about the nature of nonentrenched concepts, these experiments also required subjects to solve a set of novel reasoning problems, which were of varying degrees of difficulty. They therefore provided a way to assess the extent to which intelligence is associated with the ability to reason within new conceptual systems.

These experiments presented an opportunity to compare different structural models for nonentrenched concepts. The potential effects of linguistic familiarity and conceptual naturalness on nonentrenchment can be described in terms of five basic models. Analysis-of-variance contrast weights are used to represent patterns of reaction time and error rates that characterize each of these models.

Model 0

In Model 0, the null case, there is no effect for either linguistic unfamiliarity or conceptual unnaturalness. This model implies that there is no psychological reality in the nonentrenchment construct. In essence, this model represents the null hypothesis.

Model 1

In Model 1, the locus of nonentrenchment can be found entirely in conceptual naturalness; according to this formulation, linguistic unfamiliarity does not contribute to nonentrenchment.

Model 2

Model 2 shows the complementary situation, in which the locus of nonentrenchment can be found entirely in linguistic unfamiliarity; according to this formulation, conceptual unnaturalness does not contribute to nonentrenchment.

Model 3

Model 3 is essentially an extension of Models 1 and 2, in that it describes the situation in which linguistic unfamiliarity and conceptual unnaturalness are both important, such that their effects are additive. Model 3 distinguishes between two levels of nonentrenchment. On one level, nonentrenchment is characterized by using either familiar names to denote unnatural occurrences or unfamiliar names to denote natural occurrences--at this level, there is no predicted difference in difficulty between these forms of nonentrenchment. At a second level, there is a more difficult form of nonentrenchment that involves using unfamiliar names to denote unnatural occurrences.

Model 4

Finally, Model 4 describes the situation in which nonentrenchment is defined by an interaction between linguistic unfamiliarity and conceptual unnaturalness. According to this model, there are two types of nonentrenched concepts, one in which familiar names denote unnatural occurrences, and another in which unfamiliar names denote natural occurrences. The locus of difficulty is in pairing the familiar (either concepts or language) with the unfamiliar (again, either

concepts or language).

In a first experiment, subjects were 96 Yale undergraduates who participated for course credit, monetary payment, or both. Subjects were randomly assigned to one of four conditions, with 24 subjects in each condition. A separate group of 25 undergraduates taking a developmental psychology course at Yale gave responses to a set of background survey questions.

The basic materials were selection-task items presented via a tachistoscope and psychometric inductive and deductive ability tests presented in a paper-and-pencil format. In addition to the standardized ability tests, subjects were given a set of problems that have previously been used in the study of insight (Sternberg & Davidson, 1982).

In the selection task, items were modeled on previous problems used by Sternberg (1982). Items were developed in which a common "deep" structure was used to generate four different "surface" structures. In each of the four sets of items, the problems were based on the initial premise that the seasons of the year allow one to predict certain occurrences in nature and that, in turn, these occurrences identify what a given season is. Each problem contained two pieces of information. The first piece of information described a situation at the beginning of a season and the second piece of information provided "follow-up" data from the end of the same season. Because each of the four variations of the task were similar, only one version will be described in detail. The other variations will be described more briefly.

The premise that served as the model for the other three versions of the task stated that in New Haven, the beginning and end of each season are marked by the facts that the leaves will be either green or brown. In summer, the leaves are green at the beginning and at the end of the season. In fall, the leaves are green at the beginning of the season but are brown at the end. In winter, the leaves are brown at both the beginning and end of the season. And, in spring, the leaves are brown at the beginning of the season but are green at the end. Subjects were required to use this information to solve a series of reasoning problems.

Each problem was presented on one card. Each term of a problem could contain one of two forms of information. The description could be either a picture of the leaves, indicated by a green or brown circle, or the name of a season that represents a decision about what season it is, based on the color of the leaves at the time the observation was made. Information about the leaves at the beginning of the season appeared on the left and information about the leaves at the end of the season appeared on the right. Because each of two descriptions of the leaves (one at the beginning of the season and one at the end) could take either of two physical forms (brown or green) or four verbal forms (an inference based on a season name), there were 6 x 6, or 36, distinct items. A complete listing of the problems can be found in the original paper.

The subject's task was to describe the leaves at the end of the season, based on the information provided in the problem. If the given description for the end of the season was a name, the subject had to indicate the correct color of the leaves. There were always three choices, from which the subject had to choose the correct one. These alternatives appeared below the problem stem.

There were four different types of problems. Items either had two season names, a picture followed by a season name, a season name followed by a picture, or two pictures. In the first two types of problems, subjects had to determine the color of the leaves at the end of the season. In the other two problems, subjects had to give the name of the season consistent with the given information. Subjects

were alerted to a further complexity in the selection task, which also applies in the real world. At the beginning of a season, it is impossible to distinguish summer from fall or spring from winter, if the only available information is the initial color of the leaves. Also, the names "summer" and "winter" imply that the leaves will remain the same color, whereas the names "spring" and "fall" imply that the leaves will change color by the end of the season. For problems in which the first term was the name of the season, this name correctly described the color of the leaves at the beginning of the season, but only predicted what color the leaves would be at the end. This prediction might not correspond to the color that was described by the second term. Thus, it was not possible to know for certain the color of the leaves at the end of the season or to know for certain what the true season was. When the first term of the problem was a picture of the leaves, this complexity did not exist, because a physical description carries no implication regarding the future physical appearance of the leaves.

Although this uncertainty in prediction did not exist for information describing the leaves at the end of the season, there was a related problem associated with the second term. When a season name described the leaves at the end of the season, it could be assumed to provide correct information about both the beginning and ending color of the leaves, because assessments of season made late in the season were based on observations of the leaves throughout the entire season. For the problems in which the second term was a name, however, this season name could be "inconsistent" with the starting color of the leaves, as defined in the first term of the problem. For example, if the first term was "summer," this name means that the leaves were green at the beginning of the season and predicts that they would be green at the end. If the second term was "spring," this name means that the leaves were brown at the start of the season and eventually turned green. Because the leaves cannot be both brown and green at the beginning of the season, this problem describes an inconsistent situation; as a result, it was impossible to determine the color of the leaves at the end of the season. The correct answer was thus "inconsistent."

To summarize, physical descriptions, which carried no necessarily correct implication for what the leaves would look like at another time, were always accurate with respect to the appearance of the leaves at the time of the description. However, they might not be accurate with respect to the appearance of the leaves at the end of the season.

This experiment attempted to assess the extent to which various conceptual systems are more or less "entrenched" by comparing how different problem contents and forms affect reasoning. In the form mentioned above, subjects were required to reason within an entrenched framework. The other three forms of this task varied either the "naturalness" of the concepts or the type of language used, or both. It is expected that leaves will alternate between green and brown as the seasons change. However, it is not at all normal to expect that rocks will change from orange to blue with the passage of seasons. Similarly, the terms summer, fall, winter, and spring carry certain connotations about the seasons they name, but the neologisms soob, trit, blen, and mave do not carry any unequivocal information about the physical world. Because there are two types of concepts (natural and unnatural) and two types of names (familiar and novel), there are 2 x 2 or four possible versions of the season - color information.

In a second condition, subjects were told about the distant country of Latzania, where the leaves change color just as they do in New Haven, but the seasons are called trit, blen, mave, and soob. In a third condition, subjects were told about the planet Kyron, where the seasons are called summer, fall, winter, and

spring, but are marked by the fact that rocks change from blue to orange or from orange to blue. In the fourth condition, subjects were told about the planet Kyron, where the seasons are called trit, blen, mave, and soob, and can be distinguished by the fact that the rocks vary from orange to blue, according to a systematic pattern.

Using each of the conditions described above, three more sets of 36 problems were generated that were structurally identical to those described for the case of New Haven. The only difference among the four tasks involved the extent to which the conceptual system was "entrenched" as defined by a particular concept-language combination. This manipulation of content made it possible to identify the locus of "nonentrenchment."

The ability tests that were used included geometric series completion (Abstract Reasoning) from the Differential Aptitude Test, letter and number series completions (Reasoning) from the SRA Primary Mental Abilities, adult level, and deductive syllogisms and confirming the valdity of conclusions (subtests 9) from the Watson-Glaser Critical Thinking Appraisal.

The survey requested subjects to rate how common it is, in their experience, for either leaves or rocks to change colors with the seasons. In addition, subjects were given an array with the names "summer," "fall," "winter," and "spring" on one side and the four possible blue-orange or brown-green pairings on the other side. Their task was to match one season name with one of the four physical occurrences. Subjects answered questions about either leaves or rocks, but not both. The survey validated our notions of what was "nonentrenched."

The overall design placed subjects within a 2 x 2 between-subjects factorial arrangement. The primary dependent variable was solution latency; a secondary dependent variable was error rate. In item construction, independent variables were the six possible state descriptions (e.g., in the case of New Haven, summer, fall, winter, spring, G (green circle), and B (brown circle) crossed with the two possible times of occurrence (the beginning and end of a season)). All subjects saw each of the 36 possible item types three times, with the correct option in a different location each time. The items in each task variant were grouped into three blocks so that a subject had to complete an entire set of 36 items before seeing an item for a subsequent time.

Each subject was randomly assigned to one of the four versions of the selection task. For items in which subjects had to determine the color of the leaves at the end of the season, distractors consisted of an incorrect picture and an "indeterminate" (I) option, to correspond to the possibility that the information in the problem could be describing a self-contradictory and hence, indeterminate situation. For the items that required subjects to determine the name of the season, distractors consisted of two of the three possible word distractors balanced over the three replications of the task for a given experiment. Thus, each possible word distractor appeared equally often across item replications. In the three versions that described the seasons in Latzania or on Kyron, four counterbalanced forms of the items were constructed so that each season name was paired with each of the four physical occurrences (two concepts describing a physical change and two concepts describing constant physical states) only once. For example, in Latzania, the season that would correspond to summer was soob in Form A, trit in Form B, blen in Form C, and mave in Form D. The scheme was followed for each of the other three seasons. This method of counterbalancing was not applied to the New Haven scenario, however, because this scenario did not have any possible alternative forms. Subjects who filled out the survey were given one of two alternative forms of the questions. The forms differered in the order that the season

names were listed, so that subjects would not be biased in favor of choosing a particular season name for any one color change.

Selection-task items were administered on an Iconix tachistoscope with attached millisecond timer. Subjects signaled their responses by pressing the button corresponding to the appropriate answer option. Psychometric ability tests, insight problems, and the survey were all administered in written form.

The instructions for the experiment were rather lengthy and required the subjects to learn how to solve an entirely new set of reasoning problems. Because of the complexity of the task, after each subject finished reading the instructions, the experimenter reviewed the essential elements involved in each of the four types of problems. Then subjects received eight practice items, two of each of the four types of problems described earlier. When needed, extra practice items were provided until subjects were able to give correct responses and to demonstrate that they were aware of the different requirements of each problem. Subjects were instructed to solve the items as rapidly as they could under the constraint that they be as accurate as possible. After the practice trials were over, subjects received three randomized blocks of 36 problems, each of a different type. Each of the 36 problem types appeared once in each of the three blocks. Items were drawn on separate 6 x 9 - inch cards. Each answer option was correct equally often in each block (12 times per block). The experimenter initiated each trial. The millisecond clock started as soon as the subject pressed one of the three answer buttons. In general, feedback was not provided during selection-task trials, unless subjects made three errors in a row. This feedback was given to ensure that subjects were aware of the various intricacies involved in the different types of problems. Of the 96 subjects who participated in this experiment, 12 were given feedback. The selection task usually took about one hour to administer.

The ability tests were administered at a later time in small groups. All tests were timed and subjects were told to complete the tests as quickly and accurately as possible. The various tests were always presented in the following order, under the specified time constraints: (1) letter series, (2) deductive syllogisms, (3) number series, (4) confirming the validity of conclusions, (5) abstract reasoning, and (6) insight problems.

The critical finding in this study was that the data supported Model 4: The mean reaction times were 3.83 seconds for group 1, 4.18 seconds for group 2, 4.80 seconds for group 3, and 3.76 seconds for group 4. Error rates showed the same pattern as the solution latencies. In other words, the locus of the nonentrenchment effect was not in linguistic difficulties per se, nor was it in conceptual difficulty per se, but it was in the integration of linguistic with conceptual information. In other words, subjects found it difficult to process the conceptual-projection items when novel concepts were paired with familiar linguistic tokens, or when novel linguistic tokens were paired with familiar concepts. They did not find difficult problems in which the linguistic tokens and the concepts underlying these tokens were either both novel or both nonnovel. An information-processing model accounted for over 90% of the variance in the latency data.

A second experiment was designed to replicate and generalize the findings of the first study. The major differences involved the content of the conceptual system and the words used to describe this content. Subjects were 80 Yale undergraduates. They were randomly assigned to one of four conditions, with 20 subjects in each condition. The same ability tests and insight problems were used as in the first experiment. The selection-task items used in this experiment were structurally identical to those used previously, differing only in content. The premise that serves as the model for the other three versions of the task stated that in New Haven, the beginning and end of each period of the day is marked by

the sun starting out below the horizon but eventually ending up above the horizon. In daytime, the sun is above the horizon at the beginning and at the end of this period. At dusk, the sun starts out above the horizon but eventually ends up below the horizon. And in nighttime, the sun is below the horizon at the beginning and at the end of this period.

In order to parallel the first experiment, the period of the day had to be characterized by an "unnatural" occurrence. Because it is not at all normal for the periods of the day to be characterized by minerals that change shape from oval to rectangular, this occurrence served as the unnatural concept. In one variation, subjects were told about the Hafo Indians of Western Canada, who use the names stobe, kovit, bren, and trofar to describe the daily patterns of changes in the position of the sun. In another version, subjects were told about the planet Kyron, where the periods of the day are called dawn, daytime, dusk, and nighttime, but are marked by the fact that a certain type of mineral changes from rectangular to oval shape or from oval shape to rectangular shape. In the final variation, subjects were told about the planet Kyron, where the periods of the day are called stobe, kovit, bren, and trofar, and can be distinguished by the fact that a certain kind of mineral changes shape from rectangular to oval, according to a systematic pattern. All stimuli were presented on a CRT screen.

The results of this experiment replicated the results of the preceding experiment. Once again, it was found that the locus of nonentrenchment lies in mapping unnatural occurrences onto familiar names or natural occurrences onto novel names. These data suggest that people have difficulty reasoning with new information because it is both similar to and different from the knowledge they already possess. The entrenchment construct shows that prior knowledge can both facilitate and impede our attempts to understand the world.

<u>Causal Inference with Verbal versus Abstract Materials</u>

Most of the experiments described above have required various kinds of reasoning by subjects exposed to verbal materials. To what extent are the reasoning processes exhibited in inductive problems specifically tied to verbal materials, and to what extent are they content-general, applying to abstract as well as to verbal materials? A pair of experiments by Downing, Sternberg, and Ross (1985) addressed this question. These experiments looked particularly at multicausal inference with verbal versus abstract-symbolic materials.

Consider a typical verbal multicausal-inference problem:

In City 1,
  Annual health inspection of food-service workers was stopped.
  A new type of pesticide was tried by local vegetable farmers.
    An epidemic of Hammonds disease was reported.
In City 2,
  A new type of pesticide was tried by local vegetable farmers.
  A new type of hair dye was used in the area.
    An epidemic of Hammonds disease was reported.
In City 3,
  A new type of pesticide was tried by local vegetable farmers.
  There was a water-main break.
    An epidemic of Hammonds disease was reported.
In City 4,
  Annual health inspection of food-service workers was not stopped.
  A new type of hair dye was not used in the area.
    An epidemic of Hammonds disease was not reported.
In City 5,

There was a water-main break.
Annual health inspection of food-service workers was not stopped.
An epidemic of Hammonds disease was not reported.
In another city,
Annual health inspection of food-service workers was stopped.
A new type of pesticide was tried by local vegetable farmers.
What is the likelihood that an epidemic of Hammonds disease would occur?

This same problem could be presented in abstract format, with letters replacing sentences in the problem.

In order to understand the experiments, it is necessary first to understand some alternative models of multicausal inference.

Models of Multicausal Inference Based on Logical Evidence Types

The mean model. We expected people to consider all of the events, or possibly causal factors, in the informational situations they were given on which to base their multicausal assessments. Furthermore, we expected them to combine the information about different types of causal evidence for the various events in the hypothetical situation they were called on to evaluate, by averaging the unicausal likelihoods that each event in the hypothetical situation would lead to the outcome, or possible consequence, of concern. We called this type of strategy for multicausal inference the mean model, and modeled the multicausal estimates that would result from its use with the mean frequency across events of each of the four logical evidence types: confirmation by joint presence, violation of necessity, violation of sufficiency, and confirmation by joint absence. This model is of course mathematically equivalent to one in which people first find the unicausal likelihood for each of the events in the hypothetical situation, and then average these unicausal likelihoods to find the multicausal likelihood (see Schustack & Sternberg, 1981). In our present experiments, we could not distinguish between these two process models of multicausal inference.

The sum model. In contrast, a second approach to multicausal evaluation would be to consider information for all of the events in the hypothetical situation, as before, but to combine the information for the different events in such a way that the strength of each evidence type would be affected by the number of events in the situation. We investigated this type of strategy using the sum model, in which we modeled people's causal evaluations on the basis of the frequencies of the four logical evidence types summed across events. Using this approach, someone asked to evaluate the problem above would find the frequency of each of the evidence types in past situations, and sum these frequencies across these events to get the strength of each evidence type for the situation. Note that this model is equivalent to finding the unicausal likelihood for each of the events, and then taking one of the unicausal likelihoods and augmenting it according to some function of the unicausal likelihoods for the other events. According to the sum model, adding an event to a situation could never decrease the multicausal estimate, whereas according to the mean model, it could.

The max model. A third possibility would be for people to weigh only information for the one event in the hypothetical situation that was seen as most likely to cause the outcome of concern in making multicausal assessments. This type of approach is consistent with models of misguided parsimony and many findings suggesting that people consider only one of a number of possible causes in making causal inferences. We modeled the causal evaluations that people would give using this approach with the frequencies of the four logical evidence types for the single most llikey causal event where causal strength was determined by

subtracting the amount of disconfirming evidence from the amount of confirming evidence (cf. Schustack & Sternbrg, 1981). We called this approach the max model because it postulated that people consider only the one event in the hypothetical situation with the maximum strength. Someone using this approach to evaluate the problem above would find the frequencies of the four evidence types in order to determine which was most likely to lead to the outcome, and use the strengths of the four evidence types for that event only. Note that in the max model, the multicausal likelihood for a set of events is the same as the unicausal likelihood of the most causal event in the set.

The subset model. The fourth and final possibility that we considered is that when people select events to consider in making causal inferences, they are biased by a type of halo effect, such that they selectively weight only those events that are consistent with their *overall* impression about whether or not the outcome is likely to occur. If the overall impression of the hypothetical situation being evaluated were that it favored the outcome, only information from events that also favored the outcome would be considered in making the likelihood estimates. On the other hand, if the overall impression were unfavorable, only those events that were also unfavorable would be considered. This approach was modeled using the subset model. It postulated that the strengths of the evidence types for the situation as a whole were determined by the mean frequencies of the four logical evidence types for that subset of events in the hypothetical situation that had *causal strengths of the same sign as the overall causal strength for the situation.*

Someone using this approach to evaluate the problem above would gather information about the frequency of the four logical evidence types for each of the four events, and use this information to determine the overall causal strength of the situation (by subtracting the frequency of disconfirming evidence from the frequency for confirming evidence). Note that in this model, people are averaging the unicausal likelihoods for those events consistent with their overall impression about the outcome in order to give estimates of multicausal likelihood. Also note that when the overall strength o! the situaton was neither positive nor negative, *indicating that the outcome was neither likely nor unlikely to occur,* mean frequencies were taken across events. In such cases, where a strong overall impression about the outcome was essentially lacking, people would have no basis for discounting one set of events in favor of another.

Additional Types of Causal Evidence

Representativeness. In evaluating the likelihood that a set of events would lead to a particular outcome, knowledge that a previous situation exactly resembling this hypothetical situation had resulted in the outcome would influence people to increase their estimates of multicausal likelihood. In contrast, *knowledge of an exactly matching previous situation that had not resulted in the* outcome would influence people to decrease their estimates of likelihood. For example, someone using the representativeness heuristic would consider events more likely that had led to a given outcome before.

We thought that people might also be biased by previous situations resembling the hypothetical situation in that they were exactly the opposite to the hypothetical situation (events present in the hypothetical situation would all be absent in such previous situations). If the outcome of a previous situation exactly opposite to the hypothetical situation had been present, people would decrease their estimates of multicausal likelihood. If the outcome in such a situation had been absent, on the other hand, they would increase their estimates of multicausal likelihood.

The degree to which the hypothetical situation was representative of causal models suggested by the informational situations was indexed by the following two variables:

1. Positive representativeness. This variable was given a value of 1 if the entire set of events in the hypothetical situation had occurred in one or more of the information situations and the outcome of these situations was positive. It was given a 0 if the entire set of events did not occur in any one informational situation or the set of events occurred in more than one informational situation, but these situations had conflicting outcomes. A value of -1 was assigned if the entire set of events occurred in one or more informational situations and the outcome of all of these situations was negative.

2. Negative representativeness. A value of 1 was used to indicate that the entire set of events was absent in one or more informational situations where the otucome did not occur. A value of 0 was used to indicate that the entire set was not absent in any one situation or that it was absent in more than one situation, but that the outcomes of these situations were conflicting. The variable was given a value of -1 when the entire set of events was absent in one or more informational situations where the outcome was positive.

The strength of alternative causes. Finally, we expected people to be influenced by the causal strength of events that were not among those in the situation they were to evaluate. In particular, we would expect people to temper their estimates of the likelihood that a set of events would lead to an outcome in a new situation to the extent that there were other events that appeared to be strongly related to the outcome in previous situations. Although it is possible that different sets of events could each be sufficient to lead to a particular outcome, people generally discount the causal role of possible factors when there are other possible factors that seem very strong (Schustack & Sternberg, 1981).

3. Strength of alternative causes. The causal strength of each event described in the informational situations, but not in the hypothetical situation, was calculated by subtracting the amount of disconfirming evidence (the sum of violation of necessity and violation of sufficiency) from the amount of confirming evidence (the sum of confirmation by joint presence and confirmation by joint absence) for that event. The causal strength of the strongest of these alternative causes was then used as an index of the strength of the alternative explanations for the outcome.

Experiments

Two experiments were conducted in order to determine how subjects solve multicausal-inference problems. In the first experiment, problems were presented in abstract format (with letters as events). In the second experiment, problems were presented in both abstract and verbal format (with sentences as events, as per the example). Subjects in the first experiment were 47 students at Yale University. Subjects in the second experiment were 74 Yale undergraduates, 34 of whom were in the abstract condition and 40 of whom were in the verbal condition. In each experiment, subjects received 60 causal-inference problems. In the second experiment, the abstract and verbal problems were isomorphic, that is, differed only with respect to content, not with respect to form. Multiple-regression modeling of probability judgments was used to determine how subjects solved the problems.

As the results of the previous experiments suggest, the strategy found by Schustack and Sternberg (1981) to account for unicausal inference appears to form the foundation for multicausal inference as well. The same four logical evidence types--confirmation by joint presence, violation of necessity, violation of

sufficiency, and confirmation by joint absence--account for most of the variance in likelihood ratings for multicausal as well as unicausal inference problems. In unicausal inference, people simply keep track of the frequency of these evidence types for the one event under consideration. Then, by using a generally additive rule, they increase their estimates of unicausal likelihood as the frequency of confirming evidence increases, and decrease them as the frequency of disconfirming evidence increases. In genral, people place the greatest emphasis on evidence concerning sufficiency, and comparatively little on evidence concerning necessity. In multicausal inference, it seems that people may first determine the unicausal likelihood of each of the events under consideration, and then use these likelihoods to determine the multicausal likelihood of the situation being evaluated.

One aspect in which the unicausal assessments made during multicausal inference differ from those made during strictly unicausal inference is in the consideration that is given to the strength of alternative causes. When the task requires that people assess only one event, they appear to take the causal strength of other events into consideration in their estimations of likelihood, decreasing their estimates of unicausal likelihood for a particular event to the extent that there are other events that appear to be strongly related to the outcome. However, when the task requires many unicausal assessments to be made and compared or integrated, individual unicausal assessments no longer appear to take the strength of alternative explanations into acocunt. Because there is much more information to be considered during multicausal inferene than during unicausal inference, it is not surprising that people making multicausal inferences tend to ignore the relationships of other events to the outcome of concern. People seem to behave as if these other factors were totally irrelevant.

In arriving at their final estimates of multicausal likelihood, people also took into account the representativeness of the set of events under consideration, thereby using a strategy not relevant for strictly unicausal inference. They used a version of the representativeness heuristic that influenced multicausal likelihoods whenever the situation under evaluation was representative of a causal model for the outcome of concern or for the absence of that outcome. When the situation under consideration was representative of a causal model for the outcome (in that it exactly matched the set of events present in a previous situation where the outcome occurred), the multicausal estimate based on the logical evidence types was increased, and when the situation was representative of a causal model for the absence of the outcome (in that it exactly matched a situation where the outcome had previously been absent), the multicausal estimate was decreased. Although reliance on this tool may result in normative decision making, it is not clear that heavy reliance on the representativeness heuristic is normative in this type of problem, where the causal model is suggested by one or at most two previous experiences with situations related to the outcome, and the informational situations have not necessarily been selected for presentation because they are good causal models.

In our studies, people differentially weighted the four logical evidence types, showing a strong tendency to favor information about joint presence over information about mixed presence and absence, and, in general, ignoring information about joint absence. The evidence types that were heavily weighted in our study all involved situations where the events of concern were present, whereas those that were not heavily weighted involved situations where the events were not present. For confirmation by joint presence, violation of sufficiency, and positive representativeness, the events were always present, but for confirmation by joint absence, violation of necessity, and negative representativeness, the events were

absent. This tendency to underweight information presented in a negative form is also consistent with prior literature.

When people were asked to evaluate multicausal inference problems presented in concrete terms, the strategies they used were different from those they used to evaluate multicausal inference problems problems presented in abstract terms in several respects.

Representativeness. People gave less weight to positive representativeness when problems were presented in concrete terms than when the same problems were presented in abstract terms. One explanation for this difference is that positive representativeness was used more in the abstract condition simply because the matches between informational and hypothetical situations were more perceptually salient than the corresponding matches in the concrete condition. Alternatively, these lower weightings may have resulted from the use of other strategies for causal reasoning, perhaps occasioned by the availability of world knowledge as a guide for problem solving. It may be that the more familiar content of the concrete problems encouraged people to draw from their experiences in solving real-world causal-inference problems and thus led them to less heavily weight those causal models that were suggested merely by one or two previous experiences. Reliance on real-world experience may also have led them to assume that the informational situations they were presented with were less likely to have been selected as possible causal models. In addition, causal models suggested by previous situations may not have been relied on heavily because their places were usurped by a priori causal models that became available. The hazards and even the fictitious epidemics described in the concrete problesm may have been sufficient to engage general causal models that had already been acquired for different types of diseases. When these a priori models were engaged, they might have blocked the use of more ad hoc models or reduced their credibility. Under such circumstances, the only apparent effect of the use of a priori models would have been the reduction of weightings for the more ad hoc causal models.

The max model and the mean model. More people made use of the max model and fewer made use of the mean model in evaluating problems in concrete form than in evaluating problems in abstract form. One explanation for the increased use of the max model in the concrete condition is that increased processing demands in the evaluation of the concrete problems may have kept people from integrating information from all of the events in the situation to be evaluated, although they did evaluate all of the events at the unicausal level. Another possibility is that people using the max model simply adopted a very conservative strategy in their causal reasoning, and decided to treat all of the problems as though effects of the separate factors were overlapping or confounded. In both conditions, people tended to use averaging rather than adding as a means for combining information.

Context effects. The effect of more familiar context on the appropriateness of strategies for causal reasoning suggests a general trend in causal reasoning, and in other types of reasoning as well. In unfamiliar situations, people may begin reasoning with the most general strategy they have acquired for that class of problems, and work their way toward more appropriate strategies only as processing limitations allow and context demands. As the reasoning task becomes more meaningful, the failings of the more general strategies would become more apparent and if processing limitations allowed awareness of these failings, subjects might be prompted to look for other strategies. This suggestion is consistent with the trend for people to abandon the more general-purpose averaging strategy when they are faced with a variety of cues in the concrete condition. In sum, then, the

use of verbal content did matter for causal reasoning, probably because people bring world knowledge to bear upon concrete problems, whether or not it is clearly relevant to the problems at hand.

Toward an Integration of Top-Down and Bottom-Up Approaches to Verbal Intelligence

All of the experiments described above have used an essentially "top-down" strategy for understanding the components of verbal intelligence. This approach is in contrast to that of Hunt, Lunneborg, and Lewis (1975), who used a bottom-up approach. In this last experiment to be described, we sought to integrate these two approaches by studying information processing in verbal intelligence that requires varying levels of processing, ranging from bottom-up to top-down (Caruso & Sternberg, 1985).

In any cognitive task, there are stimuli to be encoded, where encoding refers to the processes operating on a stimulus to allow it to be further processed. The encoded representation is then compared to another stimulus (which may be physically present, stored in working memory, or retrieved from long-term memory). Finally, a decision is made as to whether the two stimuli "match" one another, based upon instructions given to the subject. This general model can be applied to a wide range of cognitive tasks employed in individual-differences research.

Contents of cognitive tasks vary, however. Processes operate on different inputs, and, according to the model, these inputs are considered to fall along a levels-of-complexity dimension. We propose that the levels-of-complexity dimension represents a continuum of levels with the ordering of levels a function of the complexity of the content of a given task. Complexity is defined in terms of the intricacy of the representation of the task content. One can represent stimuli either in terms of lists of features of attributes of the stimuli, or by relying upon representational systems of declarative knowledge, such as the one proposed by Anderson (1976). Such systems represent information as concept nodes linked together by logical relations. In the present model, nodes represent patterns, letters, and words, and links represent logical relations between nodes. The greater the number of nodes and links between nodes that are necessary for the representation of the content of a task, the higher is the level of complexity for the task. For instance, the representation of a pattern-matching task includes physical feature nodes that are linked spatially to form a given pattern. If semantic meaning is to be extracted, then additional nodes and links specify how the physical features are linked to form words, and the words are linked to conceptual information in terms of additional nodes and links. The level of complexity of a given task is further defined as the complexity of representation necessary for completion of the task.

The levels are ordered on a continuum rather than in discrete steps. Although it is possible to specify a circumscribed number of levels of complexity, within a given "level" there would be an almost infinite number of sub-levels, or levels of difficulty. For example, at the level of complex relationships in verbal analogies, there would probably be as many levels of difficulty as there are combinations of pairs of words, or groups of words. Even at the level of letter-matching based on physical patterns, certain pairs of letters are more discriminable than others.

At a theoretical level, the model provides a framework for intelligence and information-processing research in terms of the processes involved in task performance. Methodologically, intelligence research has primarily been divided into the two camps of cognitive correlates and cognitive components. In light of the levels-of-complexity approach, we view this dichotomy as artificial, and

hypothesize that correlates researchers have been examining the same processes as have components researchers; but the level of complexity at which these processes operate differs between the two approaches.

The present study assessed the utility of the process-content model by selecting cognitive tasks from among the pool of tasks commonly employed by both correlates and components researchers. Although the model could be tested by examining the results from previous research, it is difficult to evaluate and integrate the results of previous research due to the varied methods employed in different experiments. Therefore, each of the tasks employed an identical presentation format. Performance of subjects on the cognitive tasks was examined in order to determine whether the model accurately ranked tasks along the levels-of-complexity continuum. Tasks were also correlated with verbal ability in order to determine what level of task complexity best predicted psychometrically-measured ability. Finally, the process-content model was examined by correlating the processes--encoding, comparison, and decision--at each level of complexity, with verbal ability. In other words, the present approach allowed for an analysis of both process and content.

Fifty subjects were recruited from local colleges and universities. In order to obtain a stable estimate of verbal ability, subjects received the Verbal Reasoning subtest of the Differential Aptitudes Test, the Nelson-Denny Reading Test, and a vocabulary test from the French Kit of Reference Tests for Cognitive Factors (in addition to the vocabulary section of the Nelson-Denny).

Subjects received a set of experimental tasks via a Northstar Horizon microcomputer. The tasks were all presented in "precued" format, so that a first part of each item was presented until subjects pressed a ready button, and then a second part of each item was presented until subjects pressed a response button.

The seven experimental tasks used in the experiment were (a) the Posner letter-matching task (Posner & Keele, 1967), (b) the visual-search task of Estes and Taylor (1964), (c) the memory-search task of Sternberg (1969), (d) a category-match task, in which subjects had to indicate whether or not a given item was a member of a given category, (e) Clark and Chase's (1972) sentence-verification task, (f) verbal analogies as used by Sternberg (1977), and (g) a synonyms task, in which subjects had to indicate which of the two words provided the best synonym for a given target.

Are the levels of complexity we proposed on a continuum, or is there a clear dichotomy between low-level correlates tasks and high-level components tasks? In order to determine the underlying structure of the complexity of the tasks, the eight task variables (with separation of Posner physical and name identity) were intercorrelated, and the correlation matrix submitted to a principal-axis factor analysis. A single factor, with an eigenvalue of 5.48, was extracted. The second factor (after varimax rotation) had an eigenvalue of only 0.70. The single factor solution suggests that there is a general cognitive ability factor underlying the tasks, rather than more specific factors representing separate levels of complexity. In addition, task modeling revealed that good to excellent fits could be obtained with mathematical models based upon our notion of encoding, comparison, and decision components of information processing.

The main point to be made is that bottom-up and top-down approaches are not discretely different and nonoverlapping in the processes they tap. Rather, they appear to tap highly related processes that differ largely in terms of level of processing of information. A complete model of verbal comprehension will deal with all levels of processing simultaneously.

Conclusions

A theory of the psychology of verbal comprehension has been presented that deals with two major aspects of comprehension: the acquisition of verbal concepts and the real-time processing of verbal concepts. The theory of real-time processing deals both with the performance components, or non-executive processes of comprehension, and with the metacomponents, or executive processes of comprehension. Thus, the theory as a whole deals with all three aspects of verbal information processing in my more general theory of intelligence (Sternberg, 1985): knowledge-acquisition components, performance components, and metacomponents.

The subtheory of knowledge-acquisition components specifies three processes--selective encoding, selective combination, and selective comparison--that are applied to context cues, but the efficacy of whose application is affected by mediating variables that make it differentially difficult to apply the processes to the cues. Multiple experiments were conducted in order both to test the subtheory and to examine its implications for the training of verbal-comprehension skills. The experiments provide support for the subtheory, and also show its efficacy as a basis for training decontextualization skills. Indeed, the results suggest that learning from context is a good method for training vocabulary acquisition skills only when the training is theoretically based.

The subtheory of performance components in real-time verbal comprehension attempts to deal with the question of how individuals process the meanings of words at the time they are encountered, for example, when an individual is taking a vocabulary test. In this subtheory, words are understood as comprising sets of defining and characteristic attributes, which are interrogated when one is making a decision about the meaning of a given word, and about how this meaning compares to that of another word. A set of experiments showed that the subtheory of verbal comprehension provides a good account of people's real-time processing of word meanings. The theory can be extended as well to processing of novel concepts and to causal inference.

The subtheory of metacomponents deals with how executive processes are used in reading. In particular, how do subjects allocate their time and mental resources so as to maximize their reading comprehension, given that both time and resources are limited? The results suggest that better and poorer readers allocate their time and resources differently, and in particular, that better readers are more planful in their time and resource allocation. Better readers, for example, do more global planning for reading than do poorer readers.

The theory of verbal comprehension presented here is obviously incomplete in the scope of questions about verbal comprehension with which it can deal: For example, it says nothing about use of phonics, grammar, and syntax in verbal comprehension. Thus, no claim can be made that the theory is comprehensive. At the same time, the theory probably covers more ground than many other extant theories of verbal comprehension, which tend to be even narrower in scope. Most importantly, the proposed theory carries us some way beyond psychometric notions, such as that of Thurstone (1938), or even more recent notions of multiple intelligences, such as that of Gardner (1983), which specify "verbal ability" or "verbal intelligence" without providing a clear and coherent anatomy of the domain. Indeed, the advantage of information-processing theories is that they can go beyond naming a domain and actually help theorists specify just what kinds of mental processing occur within the domain.

Although alternative approaches to understanding verbal comprehension have been compared in this chapter, the comparison is in no way intended to suggest

that certain approaches are *useful* and others *not* so. *To the contrary, multiple* approaches must be used in conjunction to elucidate as many aspects of verbal comprehension as possible. Thus, for example, although my own research has tended to be top-down rather than bottom-up or knowledge-based with respect to the investigation of knowledge-acquisition processes, it should be self-evident that bottom-up as well as top-down processes are involved in knowledge acquisition, and that such processes are interactive with knowledge: They are knowledge-driven processes *in search of further knowledge*. Thus, *further* development of theories of verbal comprehension will almost certainly have to utilize a knitting of current theories and approaches in order more fully to understand the phenomena that constitute the psychology of verbal comprehension.

References

Anderson, J. R. (1976). Language, memory, and thought. Hillsdale, NJ: Erlbaum.

Anderson, N. H. (1979). Algebraic rules in psychological measurement. American Scientist, 67, 555-563.

Anderson, R. C., & Freebody, P. (1979). Vocabulary knowledge. (Tech. Rep. No. 136). Champaign, IL: University of Illinois, Center for the Study of Reading.

Baddeley, A. D. (1968). A 3-minute reasoning test based on grammatical transformation. Psychonomic Science, 10, 341-342.

Bisanz, G. L., & Voss, J. F. (1981). Sources of knowledge in reading comprehension. In A. Lesgold & C. A. Perfetti (Eds.), Interactive processes in reading. Hillsdale, N. J.: Erlbaum.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. Cognitive Psychology, 3, 193-209.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1). Hillsdale, NJ: Erlbaum.

Brown, A. L. (1980). Metacognitive development and reading. In R. Spiro, B. Bruce, & W. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new "structure of intellect." In L. B. Resnick (Ed.), The nature of intelligence. Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1981). Ability and task difficulty in cognitive psychology. Educational Researcher, 10, 11-21.

Caruso, D. R., & Sternberg, R. J. (1985). *Toward an information-processing model* of intellectual ability. Manuscript submitted for publication.

Charness, N. (1981). Aging and skilled problem solving. Journal of Experimental Psychology: General, 110, 21-38.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Erlbaum.

Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257-274.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. Cognitive Psychology, 3, 472-517.

Clark, H. H., & Clark, E. V. (1977). Psychology and language. New York: Harcourt Brace Jovanovich, Inc.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. Psychological Review, 82, 407-28.

Collins, A., & Smith, E. E. (1982). Teaching the process of reading comprehension. In D. K. Detterman & R. J. Sternberg (Eds.), How and how much can intelligence be increased? Norwood, NJ: Ablex.

Cornelius, S. W., Willis, S. L., Blow, S., & Baltes, P. B. (1983). Training research in aging: Attention processes. Journal of Educational Psychology, 75, 257-270.

Curtis, M. E. Word knowledge and verbal aptitude. Unpublished manuscript, University of Pittsburgh, 1981.

Daalen-Kapteijns, Van, M. M., & Elshout-Mohr, M. (1981). The acquisition of word meanings as a cognitive learning process. Journal of Verbal Learning and Verbal Behavior, 20, 386-399.

Daneman, M. (1984). Why some people are better readers than others: A process and storage account. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 2). Hillsdale, NJ: Erlbaum.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-466.

Downing, C. J., Sternberg, R. J., & Ross, B. (1985). Multicausal inference: Evaluation of evidence in causally complex situations. Journal of Experimental Psychology: General, 114, 239-263.

Estes, W. K., & Taylor, H. A. (1964). A detection method and probabilistic models for assessing information processing from brief visual displays. Proceedings of the National Academy of Sciences, 52, 446-454.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. Resnick (Ed.), The nature of intelligence. Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), Children's oral communication skills. New York: Academic Press.

Frege, G. (1952). On sense and reference. In P. Geach & M. Black (Eds.), Translations from the philosophical writings of Gottolb Frege. Oxford: Basil, Blackwell & Mott.

Freyd, P., & Baron, J. (1982). Individual differences in acquisition of derivational morphology. Journal of Verbal Learning and Verbal Behavior, 21, 282-295.

Gardner, H. (1983). Frames of mind: the theory of multiple intelligences. New York: Basic Books.

Goldberg, R. A., Schwartz, S., & Stewart, M. (1977). Individual differences in cognitive processes. Journal of Educational Psychology, 69, 9-14.

Guilford, J. P. (1967). The nature of human intelligence. New York: McGraw Hill.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. Journal of Verbal Learning and Verbal Behavior, 18, 441-461.

Hogaboam, T. W., & Pellegrino, J. W. (1978). Hunting for individual differences: Verbal ability and semantic processing of pictures and words. Memory and Cognition, 6, 189-193.

Holyoak, K. J., Glass, A. L., & Mah, W. A. (1976). Morphological structure and semantic retrieval. Journal of Verbal Learning and Verbal Behavior, 15, 235-247.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized ability intelligences. Journal of Educational Psychology, 57, 253-270.

Hunt, E. B. (1978). Mechanics of verbal ability. Psychological Review, 85, 109-130.

Hunt, E. B. (1980). Intelligence as an information processing concept. British Journal of Psychology, 71, 449-474.

Hunt, E. B. (1984). Verbal ability. In R. J. Sternberg (Ed.), Human abilities: An information processing approach. San Francisco: Freeman.

Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? Cognitive Psychology, 7, 194-227.

Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. Journal of Experimental Psychology: General, 108, 151-181.

Jensen, A. R. (1980). Bias in mental testing. New York: Free Press.

Johnson, D. D., & Pearson, P. D. (1978). Teaching and reading vocabulary. New York: Holt.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. Psychological Review, 87, 329-354.

Kaye, D. B., & Sternberg, R. J. (1983). Development of lexical decomposition ability. Unpublished manuscript.

Keating, D. P., & Bobbitt, B. L. (1978). Individual and developmental differences in cognitive-processing components of mental ability. Child Development, 49, 155-167.

Keil, F. C. (1979). Semantic and conceptual development. Cambridge, MA: Harvard University Press.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. Psychological Review, 88, 197-227.

Keil, F. C. (1984). Mechanisms of cognitive development and the structure of knowledge. In R. J. Sternberg (Ed.), Mechanisms of cognitive development. San Francisco: Freeman.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Erlbaum.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. Psychological Review, 85, 363-394.

Lansman, M., Donaldson, G., Hunt, E., & Yantis, S. (1982). Ability factors and cognitive processes. Intelligence, 6, 347-386.

Manelis, L., & Tharp, D. A. (1977). The processing of affixed words. Memory and Cognition, 5, 690-695.

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. Child Development, 48, 986-992.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. Child Development, 50, 643-655.

Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), Children's oral communication skills. New York: Academic Press.

Marshalek, B. (1981). Trait and process aspects of vocabulary knowledge and verbal ability (NR154-376 ONR Technical Report No. 15). Stanford, CA: School of Education, Stanford University.

Masur, E. F., McIntyre, C. W., & Flavell, J. H. (1973). Developmental changes in apportionment of study time among items in a multitrial free recall task. Journal of Experimental Child Psychology, 15, 237-246.

Matarazzo, J. D. (1972). Wechsler's measurement and appraisal of adult intelligence (5th ed.). Baltimore: The Williams & Wilkins Company.

McNamara, T. P., & Sternberg, R. J. (1983). Mental models of word meaning. Journal of Verbal Learning and Verbal Behavior, 22, 449-474.

Miller, G. A., & Johnson-Laird, P. N. (1976). Language and perception. Cambridge, MA: Harvard University Press.

Murrell, G. A., & Morton, J. (1974). Word recognition and morphemic structure. Journal of Experimental Psychology, 102, 963-968.

O'Rourke, J. P. (1974). Toward a science of vocabulary development. The Hague: Mouton.

Perfetti, C. A. (1983). Individual differences in verbal processes. In R. F. Dillon & R. R. Schmeck (Eds.), Individual differences in cognition (Vol. 1). New York: Academic Press.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. Journal of Educational Psychology, 67, 461-469.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and individual differences. In P. Carpenter & M. Just (Eds.), Cognitive processes in comprehension: The 12th annual Carnegie symposium on cognition. Hillsdale, NJ: Erlbaum.

Posner, M. I., & Keele, S. W. (1967). Decay of visual information from a single letter. Science, 158, 137-139.

Posner, M. I., & Mitchell, R. F. (1967). Chronometric analysis of classification. Psychological Review, 74, 392-409.

Rieger, C. (1975). Conceptual memory. In R. C. Schank (Ed.), Conceptual information processing. Amsterdam: North-Holland.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), Cognition and categorization. Hillsdale, NJ: Erlbaum.

Rubin, D. C. (1976). The effectiveness of context before, after, and around a missing word. Perception and Psychophysics 19, 214-216.

Rubin, G. S., Becker, C. A., & Freeman, R. H. (1979). Morphological structure and its effect on visual word recognition. Journal of Verbal Learning and Verbal Behavior, 18, 757-767.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, NJ: Erlbaum.

Rumelhart, D. E., & Norman, D. A. (1975). The active structural network. In D. A. Norman & D. E. Rumelhart (Eds.), Explorations in cognition. San Francisco: Freeman.

Russell, B. (1956). On denoting. In R. C. Marsh (Ed.), Logic and knowledge. London: George Allen & Unwin.

Salatas, H., & Flavell, J. H. (1976). Behavioral and metamnemonic indicators of strategic behaviors under remember instruction in first grade. Child Development, 47, 80-89.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ: Erlbaum.

Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. Journal of Experimental Psychology: General, 110, 101-120.

Schwartz, S. P. (1977). Naming, necessity, and natural kinds. London: Cornell University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 81, 214-241.

Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 275-290.

Stanners, R. F., Neiser, J. J., & Painton, S. (1979). Memory representation for prefixed words. Journal of Verbal Learning and Verbal Behavior, 18, 733-743.

Sternberg, R. J. (1977). Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities. Hillsdale, NJ: Erlbaum.

Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. Behavioral and Brain Sciences, 3, 573-584.

Sternberg, R. J. (1981). Intelligence and nonentrenchment. Journal of Educational Psychology, 73, 1-16.

Sternberg, R. J. (1982). Natural, unnatural, and supernatural concepts. Cognitive Psychology, 14, 451-488.

Sternberg, R. J., & Davidson, J. E. (1982). The mind of the puzzler. Psychology Today, 16, June, 37-44.

Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

Sternberg, R. J., & McNamara, T. P. (1985). The representation and processing of information in real-time verbal comprehension. In S. E. Embretson (Ed.), Test design: Contributions from psychology, education, and psychometrics. New York: Academic Press.

Sternberg, R. J., & Neuse, E. (1983). Utilization of context in verbal comprehension. Manuscript submitted for publication, 1983.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. American Psychologist, 38, 878-893.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. (1983). Teaching vocabulary-building skills: A contextual approach. In A. C. Wilkson (Ed.), Classroom computers and cognitive science. New York: Academic Press.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders's method. Acta Psychologica, 30, 276-315.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. Memory and Cognition, 7, 263-272.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. Journal of Verbal Learning and Verbal Behavior, 14, 638-647.

Thurstone, L. L. (1938). Primary mental abilities. Chicago: University of Chicago.

Tetewsky, S., & Sternberg, R. J. (1985). Conceptual versus linguistic determinants of nonentrenched thinking. Manuscript submitted for publication.

Vernon, P. E. (1971). The structure of human abilities. London: Methuen.

Wagner, R. K., & Sternberg, R. J. (in press). Executive processes in reading. In B. Britton (Ed.), Executive control processes in reading. Hillsdale, NJ: Erlbaum.

Werner, H., & Kaplan, E. (1952). The acquisition of word meanings: A developmental study. Monographs of the society for research in child development (No. 51).

Wittgenstein, L. (1953). Philosophical investigations. Oxford: Basil, Blackwell & Mott.

Author Note

Preparation of this report was supported by Contract N0001483K0013 from the Office of Naval Research and Army Research Institute. I am grateful to Cathryn Downing, Daniel Kaye, Timothy McNamara, Elizabeth Neuse, Janet Powell, Brian Ross, Sheldon Tetewsky, and Richard Wagner for the collaborations in research on verbal comprehension that helped make this chapter possible.

Figure Captions

Figure 1. A hypothetical individual's *mental representation of information* according to the proposed theory. Attributes with the letter "n" adjacent to them are specified as necessary; attributes with the letter "c" are specified as characteristic (nonnecessary).

Figure 2. A hypothetical individual's build-up of a mental representation for a story about a BLUMEN (kangaroo). Attributes with the letter "n" adjacent to them are specified as necessary; attributes with the letter "c" are specified as characteristic.

Figure 3. Flow chart for real-time model of information processing during attribute comparison.

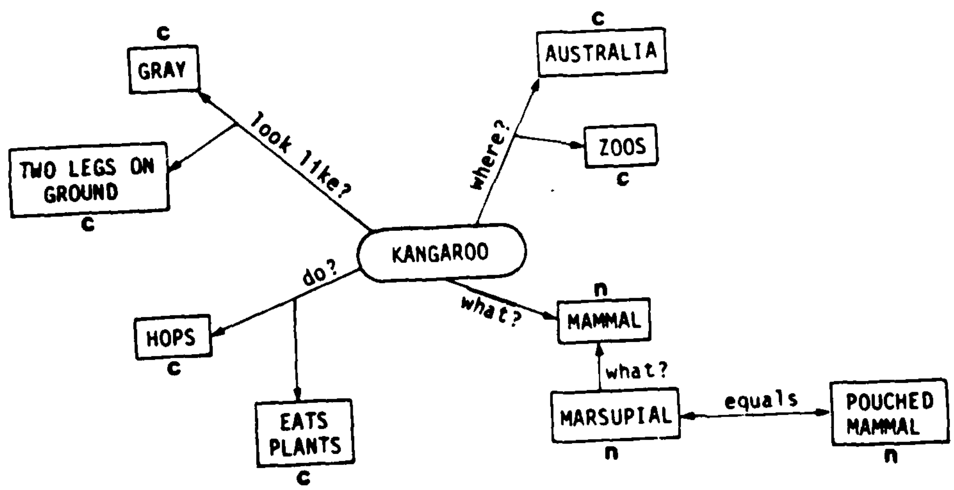

KANGAROO
look like?
GRAY
c
TWO LEGS ON GROUND
c
where?
AUSTRALIA
c
ZOOS
c
do?
HOPS
c
EATS PLANTS
c
what?
MAMMAL
n
what?
MARSUPIAL
n
equals
POUCHED MAMMAL
n

Figure 1

1. He first saw a BLUMEN during his trip to Australia.

   a. Selectively encode

      BLUMEN (saw, Australia)

      saw ↑ stative descriptive ↓ visible (n)

      Australia ↑ spatial locative ↓ in Australia (c)

   b. Selectively combine

   c. Selectively compare

      visible objects first seen in Australia
      - Aborigines
      - Animals
        - Kangaroo
        - Koala
        - Rabbit
        - Sheep
      - Plants
        - Eucalyptus

2. He had just arrived from a business trip to India and felt very tired.

   a. Selectively encode
      ∅

   b. Selectively combine
      (same as 1b)

   c. Selectively compare
      (same as 1c)

3. Looking out at the plain, he saw the BLUMEN hop across it.

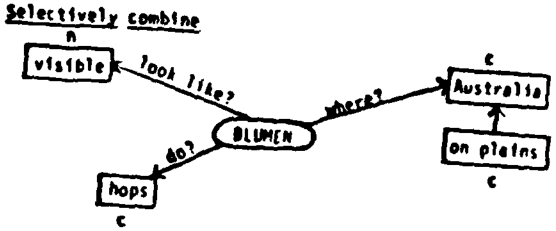
   a. Selectively encode

      BLUMEN (at the plain, hop)

      at the plain ↑ spatial ↓ on plains (c)

      hop ↑ functional descriptive ↓ hops (c)

   b. Selectively combine



   c. Selectively compare

      visible objects first seen in Australia, that hop across plains
      - Aborigines
      - Animals
        - Kangaroo
        - Koala
        - Rabbit
        - Sheep
      - Plants
        - Eucalyptus

4. It was a typical marsupial, getting its food by chewing on the surrounding plants.

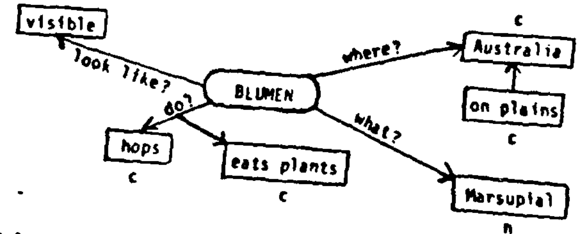
   a. Selectively encode

      BLUMEN (marsupial, chews plants)

      marsupial ↑ class membership ↓ marsupial (n)

      chews plants ↑ functional descriptive ↓ eats plants (c)

   b. Selectively combine



   c. Selectively compare

      visible objects first seen in Australia, that hop across plains, are marsupials, and chew plants

      - Animals
        - Kangaroo
        - Rabbit

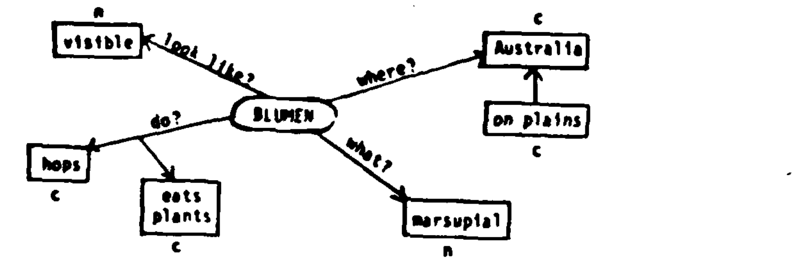
Figure 2

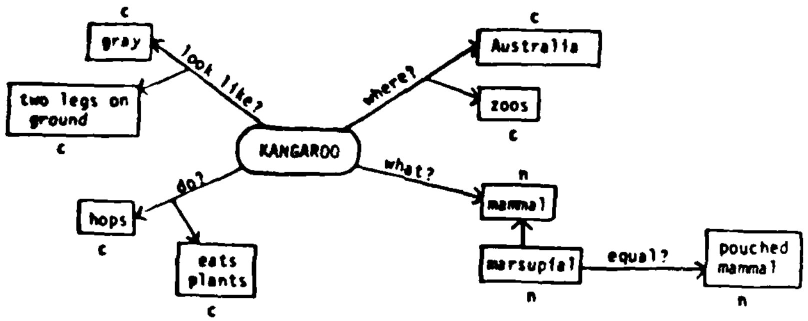
5. Squinting because of the bright sunlight and an impending headache, he noticed a young BLUMEN securely fastened in an opening in front of its mother.

a. Selectively encode

b. Selectively combine (same as 4b)

c. Selectively compare





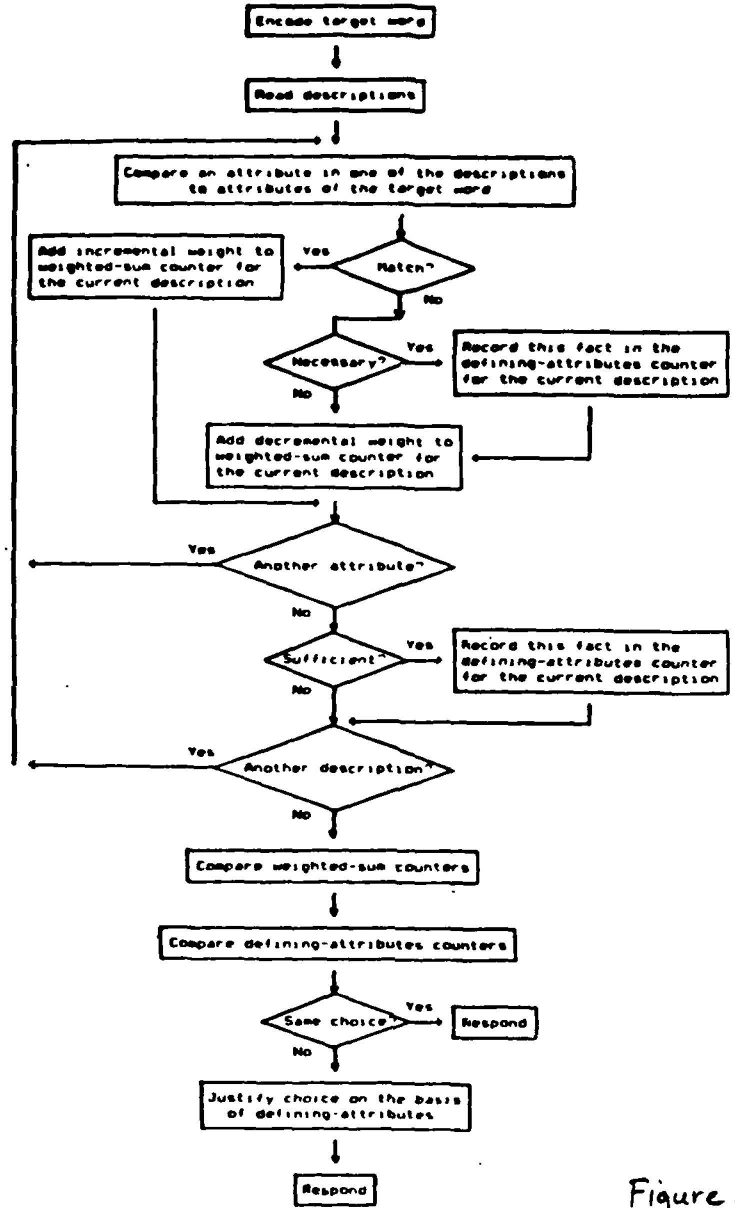


Figure 3

Publications Supported Wholly or in Part by Contract N0001483K0013 from the Office of Naval Research and the Army Research Institute

1982

Detterman, D. K., & Sternberg, R. J. (Eds.), How and how much can intelligence be increased? Norwood, N.J.: Ablex, 1982.

Sternberg, R. J. (Ed.), Advances in the psychology of human intelligence (Vol. 1). Hillsdale, N.J.: Erlbaum, 1982.

Sternberg, R. J. Casting stones: A reply to Humphreys. Journal of Educational Psychology, 1982, 74, 427-429.

Sternberg, R. J. A componential approach to intellectual development. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1). Hillsdale, N.J.: Erlbaum, 1982.

Sternberg, R. J. (Ed.), Handbook of human intelligence. New York: Cambridge University Press, 1982.

Sternberg, R. J. Lies we live by: Misapplication of tests in identifying the gifted. Gifted Child Quarterly, 1982, 26, 157-161.

Sternberg, R. J. Natural, unnatural, and supernatural concepts. Cognitive Psychology, 1982, 14, 451-488.

Sternberg, R. J. Nonentrenchment in the assessment of intellectual giftedness. Gifted Child Quarterly, 1982, 26, 63-67.

Sternberg, R. J. Reasoning, problem solving, and intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence. New York: Cambridge University Press, 1982.

Sternberg, R. J. A schism has arisen. Roeper Review, 1982, 4, 52.

Sternberg, R. J. Teaching scientific thinking to gifted children. Roeper Review, 1982, 4, 4-6.

Sternberg, R. J. Who's intelligent? Psychology Today, 1982, 16, April, 30-39.

Sternberg, R. J., & Davidson, J. E. Componential analysis and componential theory. Behavioral and Brain Sciences, 1982, 53, 352-353.

Sternberg, R. J., & Davidson, J. E. *The mind of the puzzler.* Psychology Today, 1982, 16, June, 37-44.

Sternberg, R. J., & Downing, C. J. The development of higher-order reasoning in adolescence. Child Development, 1982, 53, 209-221.

Sternberg, R. J., & Gardner, M. K. A componential interpretation of the general factor in human intelligence. In H. J. Eysenck (Ed.), A model for intelligence. Berlin: Springer-Verlag, 1982.

Sternberg, R. J., & Kaye, D. B. *Intelligence.* In H. Mitzel (Ed.), Encyclopedia of educational research (5th ed.). Washington, D.C.: American Educational Research Association, 1982.

Sternberg, R. J., & Ketron, J. L. *Selection and implementation of strategies in* reasoning by analogy. Journal of Educational Psychology, 1982, 74, 399-413.

Sternberg, R. J., Ketron, J. L., & Powell, J. S. Componential approaches to the training of intelligent performance. In D. K. Detterman & R. J. Sternberg (Eds.), How and how much can intelligence be increased? Norwood, N.J.: Ablex, 1982.

Sternberg, R. J., & Powell, J. S. Theories of intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence. New York: Cambridge University Press, 1982.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. The nature of verbal comprehension. Poetics, 1982, 11, 155-187.

Sternberg, R. J., & Salter, W. Conceptions of intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence. New York: Cambridge University Press, 1982.

Tourangeau, R., & Sternberg, R. J. Understanding and appreciating metaphors. Cognition, 1982, 11, 203-244.

1983

McNamara, T. P., & Sternberg, R. J. Mental models of word meaning. Journal of Verbal Learning and Verbal Behavior, 1983, 22, 449-474.

Sternberg, R. J. Criteria for intellectual skills training. Educational Researcher, 1983, 12, 6-12, 26.

Sternberg, R. J. Componential theory and componential analysis: Is there a "Neisser" alternative? Cognition, 1983, 15, 199-206.

Sternberg, R. J. Components of human intelligence. Cognition, 1983, 15, 1-48.

Sternberg, R. J. How much gall is too much gall? A review of Frames of mind: The theory of multiple intelligences. Contemporary Education Review, 1983, 2, 215-224.

Sternberg, R. J. Review of The Psychology of Deductive Reasoning, by J. St. B. T. Evans. British Journal of Psychology, 1983, 74, 424-426.

Sternberg, R. J. Should K come before A, B, and C? A Review of the Kaufman Assessment Battery for Children. Contemporary Education Review, 1983, 2, 199-208.

Sternberg, R. J., & Davidson, J. E. Insight in the gifted. Educational Psychologist, 1983, 18, 52-58.

Sternberg, R. J., & Gardner, M. K. *Unities in inductive reasoning.* Journal of Experimental Psychology: General, 1983, 112, 80-116.

Sternberg, R. J., & Nigro, G. Interaction and analogy in the comprehension and appreciation of metaphors. Quarterly Journal of Experimental Psychology, 1983, 35A, 17-38.

Sternberg, R. J., & Powell, J. S. Comprehending verbal comprehension. American Psychologist, 1983, 38, 878-893.

Sternberg, R. J., & Powell, J. S. The development of intelligence. In P. H. Mussen (Series Ed.) & J. Flavell & E. Markman (Volume Eds.), Handbook of child psychology (Vol. 3) (3rd ed.). New York: Wiley, 1983.

Sternberg, R. J., Powell, J. S., & Kaye, D. B. Teaching vocabulary-building skills: A contextual approach. In A. C. Wilkinson (Ed.), Classroom computers and cognitive science. New York: Academic Press, 1983.

Sternberg, R. J., & Wagner, R. K. Understanding intelligence: What's in it for educators? In A nation at risk. Washington, D.C.: National Commission on Excellence in Education, 1983.

1984

Sternberg, R. J. (Ed.). Advances in the psychology of human intelligence (Vol. 2). Hillsdale, N.J.: Erlbaum, 1984.

Sternberg, R. J. The case of the disappearing disagreements: A reply to Yussen. Developmental Review, 1984, 4, 145-147.

Sternberg, R. J. Common and uncommon issues in AI and psychology. In A. Elithorn & R. Banerji (Eds.), Artificial and human intelligence. Amsterdam: North-Holland, 1984.

Sternberg, R. J. A contextualist view of the nature of intelligence. International Journal of Psychology, 1984, 19, 307-334.

Sternberg, R. J. Does "simplicity breed content"? A reply to Jensen. Journal of Social and Biological Structures, 1984, 7, 119-123.

Sternberg, R. J. Facets of intelligence. In J. R. Anderson & S. M. Kosslyn (Eds.), Tutorials in learning and memory: Essays in honor of Gordon Bower. San Francisco: Freeman, 1984.

Sternberg, R. J. Fighting butter battles: A reply to Gardner. Phi Delta Kappan, 1984, 65, 700.

Sternberg, R. J. General intellectual ability. In R. J. Sternberg (Ed.), Human abilities: An information processing approach. San Francisco: Freeman, 1984.

Sternberg, R. J. Higher-order reasoning in post-formal-operational thought. In M. Commons & C. Armon (Eds.), Beyond formal operations: Late adolescent and adult cognitive development. New York: Praeger, 1984.

Sternberg, R. J. How can we teach intelligence? Educational Leadership, 1984, 42(1), 38-50.

Sternberg, R. J. (Ed.), Human abilities: An information-processing approach. San Francisco: Freeman, 1984.

Sternberg, R. J. If at first you don't believe, try "tri" again. Behavioral and Brain Sciences, 1983, 7, 304-315.

Sternberg, R. J. The Kaufman Assessment Battery for Children: An information-processing analysis and critique. Journal of Special Education, 1984, 18, 269-279.

Sternberg, R. J. (Ed.), Mechanisms of cognitive development. San Francisco: Freeman, 1984.

Sternberg, R. J. Mechanisms of cognitive development: A componential approach. In R. J. Sternberg (Ed.), Mechanisms of cognitive development. San Francisco: Freeman, 1984.

Sternberg, R. J. Operant analysis of problem solving: Answers to questions you probably don't want to ask. Behavioral and Brain Sciences, 1984, 7, 605.

Sternberg, R. J. Re-inventing psychology. Wilson Quarterly, 1984, 8(5), 60-71.

Sternberg, R. J. Review of H. Gardner, Frames of mind: The theory of multiple intelligences. American Scientist, 1984, 72, 394.

Sternberg, R. J. Review of M. Hunt, The Universe Within. Journal of Social and Biological Structures, 1984, 7, 85-87.

Sternberg, R. J. Testing intelligence without IQ tests. Phi Delta Kappan, 1984, 65, 694-698.

Sternberg, R. J. A theory of knowledge acquisition in the development of verbal concepts. Developmental Review, 1984, 4, 113-138.

Sternberg, R. J. Toward a triarchic theory of human intelligence. Behavioral and Brain Sciences, 1984, 7, 269-287.

Sternberg, R. J. What cognitive psychology can and cannot do for test development. In B. S. Plake & J. Mitchell (Eds.), Social and technical issues in testing: Implications for test construction and usage. Hillsdale, N.J.: Erlbaum, 1984.

Sternberg, R. J. What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. *Educational Researcher*, 1984, 13, 5-1'

Sternberg, R. J., & Lasaga, M. Approaches to human reasoning: An analytical framework. In A. Elithorn & R. Banerji (Eds.), Artificial and human intelligence. Amsterdam: North-Holland, 1984.

Wagner, R. K., & Sternberg, R. J. Alternative conceptions of intelligence and their implications for education. Review of Educational Research, 1984, 54, 197-224.

1985

Davidson, J. E., & Sternberg, R. J. Competence and performance in intellectual development. In E. Neimark, R. de Lisi, & J. H. Newman (Eds.), Moderators of competence. Hillsdale, N.J.: Erlbaum, 1985.

Downing, C. J., Sternberg, R. J., & Ross, B. Multicausal inference: Evaluation of evidence in causally complex situations. Journal of Experimental Psychology: General, 1985, 114, 239-263.

Sternberg, R. J. All's well that ends well, but it's a sad tale that begins at the end: A reply to Glaser. American Psychologist, 1985, 40, 571-573.

Sternberg, R. J. Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press, 1985.

Sternberg, R. J. The black-white differences and Spearman's g: Old wine in new bottles that still doesn't taste good. Behavioral and Brain Sciences, 1985, 8, 244.

Sternberg, R. J. Componential analysis: A recipe. In D. K. Detterman (Ed.), Current topics in human intelligence (Vol. 1). Norwood, N.J.: Ablex, 1985.

Sternberg, R. J. Controlled versus automatic processing: A reply to Fodor. Behavioral and Brain Sciences, 1985, 8, 32-33.

Sternberg, R. J. For the best on how to test: Review of Essentials of psychological testing (4th ed.), by Lee J. Cronbach. Contemporary Psychology, 1985, 30, 377-378.

Sternberg, R. J. *Implicit theories of intelligence, creativity, and wisdom.* Journal of Personality and Social Psychology, 1985, 49, 607-627.

Sternberg, R. J. Instrumental and componential approaches to the training of intelligence. In S. Chipman, J. Segal, & R. Glaser (Eds.), Thinking and learning skills: Current research and open questions (Vol. 2). Hillsdale, N.J.: Erlbaum, 1985.

Sternberg, R. J. Pretty close to ideal, anyway: A review of The Ideal Problem Solver. Human Intelligence Newsletter, 1985, 6, Spring, 10.

Sternberg, R. J. Review of Meichenbaum, Burland, Grusen, and Cameron's "Metacognitive assessment." In S. Yussen (Ed.), The growth of reflection in the *child. New York: Academic* Press, 1985.

Sternberg, R. J., & Caruso, D. Practical knowledge. In E. Eisner (Ed.), Learning the ways of knowing. Chicago: University of Chicago Press, 1985.

Sternberg, R. J., & Davidson, J. E. Cognitive development in the gifted and talented. In F. D. Horowitz & M. O'Brien (Eds.), The gifted and the talented: A developmental perspective. Washington, D.C.: American Psychological Association, 1985.

Sternberg, R. J., & McNamara, T. P. The representation and processing of information in *real-time verbal comprehension.* In S. E. Embretson (Ed.), Test design: Contributions from psychology, education, and psychometrics. New York: Academic Press, 1985.

Sternberg, R. J., & Smith, C. Social intelligence and decoding skills in nonverbal communication. Social Cognition, 1985, 2, 168-192

Wagner, R. K., & Sternberg, R. J. Practical intelligence in the real world: The role of tacit knowledge. Journal of Personality and Social Psychology, 1985, 49, 436-458.

In press

Baron, J. B., & Sternberg, R. J. (Eds.) Critical thinking: Its nature, assessment, and improvement. New York: W. H. Freeman, in press.
Berg, C. A., & Sternberg, R. J. Response to novelty: Continuity versus discontinuity in the developmental course of intelligence. In H. Reese (Ed.), Advances in child development and behavior. New York: Academic Press, in press.
Berg, C. A., & Sternberg, R. J. A triarchic theory of intellectual development during adulthood. Developmental Review, in press.
Dillon, R. F., & Sternberg, R. J. (Eds.). Cognition and curriculum design. New York: Academic Press, in press.
Marr, D. B., & Sternberg, R. J. The role of mental speed in intelligence: A triarchic perspective. In P. A. Vernon (Ed.), Speed of information processing and intelligence. Norwood, NJ: Ablex, in press.
Sternberg, R. J. (Ed.). Advances in the psychology of human intelligence (Vol. 3). Hillsdale, NJ: Erlbaum.
Sternberg, R. J. Applying componential analysis to the study of individual differences in cognitive skills. In C. Reynolds & V. Willson (Eds.), Methodological and statistical advances in the study of individual differences. New York: Plenum, in press.
Sternberg, R. J. Assessing cognitive competence: A pluralistic view. In D. A. Wilkerson & E. W. Gordon (Eds.), The development and assessment of human competence. Westport, CT: Mediax, in press.
Sternberg, R. J. Cognition and instruction: Why the marriage sometimes ends in divorce. In R. F. Dillon & R. J. Sternberg (Eds.), Cognition and curriculum design. New York: Academic Press, in press.
Sternberg, R. J. Cognitive approaches to intelligence. In B. Wolman (Ed.), Handbook of intelligence. New York: Wiley, in press.
Sternberg, R. J. Critical thinking: Its nature, measurement, and improvement. In F. Link (Ed.), Essays in cognitive education. Alexandria, VA: Association for Supervision and Curriculum Development, in press.
Sternberg, R. J. A framework for the analysis of critical-thinking skills. In J. B. Baron & R. J. Sternberg (Eds.), Critical thinking: Its nature, assessment, and improvement. New York: W. H. Freeman, in press.
Sternberg, R. J. Gee, there's more than g: A critique of Arthur Jensen's views on intelligence. In S. Modgil & C. Modgil (Eds.), Arthur Jensen: Consensus and controversy. Barcombe, England: Falmer Press, in press.
Sternberg, R. J. Human intelligence: The model is the message. Science, in press.
Sternberg, R. J. Identifying the gifted through IQ: Why a little bit of knowledge is a dangerous thing. Roeper Review, in press.
Sternberg, R. J. Implicit theories: An alternative to modeling cognition and its development. In J. Bisanz & C. Brainerd (Eds.), Formal models of cognitive development. New York: Springer-Verlag, in press.
Sternberg, R. J. Intelligence. In R. Gregory (Ed.), Oxford companion to the human mind. Oxford, England: Oxford University Press, in press.
Sternberg, R. J. Intelligence. In R. J. Sternberg & E. E. Smith (Eds.), The psychology of human thought. New York: Cambridge University Press, in press.
Sternberg, R. J. Intelligence and cognitive style. In R. E. Snow & M. J. Farr (Eds.), Aptitude, learning, and instruction (Vol. 3): Conative and affective process analyses. Hillsdale, NJ: Erlbaum, in press.

Sternberg, R. J. Intelligence applied: Understanding and increasing your intellectual skills. San Diego: Harcourt, Brace, Jovanovich, in press.
Sternberg, R. J. Intelligence is mental self-government. In R. J. Sternberg & D. K. Detterman (Eds.), What is intelligence? Contemporary viewpoints on its nature and definition. Norwood, NJ: Ablex, in press.
Sternberg, R. J. Intelligence, wisdom, and creativity: Their natures and interrelationships. In R. Linn (Ed.), Intelligence: Measurement, theory, and public policy. Champaign, IL: University of Illinois Press, in press.
Sternberg, R. J. Introduction: A framework for understanding conceptions of intelligence. In R. J. Sternberg & D. K. Detterman (Eds.), What is intelligence? Contemporary viewpoints on its nature and definition. Norwood, NJ: Ablex, in press.
Sternberg, R. J. Most vocabulary is learned from context. In M. McKeown (Ed.), The nature of vocabulary acquisition. Hillsdale, NJ: Erlbaum, in press.
Sternberg, R. J. A perspective on the prospects for Human Intelligence: Perspectives and prospects, by R. Kail & J. W. Pellegrino. Contemporary Psychology, in press.
Sternberg, R. J. Practical intelligence: Nature and origins of competence in the everyday world. In R. J. Sternberg & R. K. Wagner (Eds.), Practical intelligence: Its nature, origins, and assessment. New York: Cambridge University Press.
Sternberg, R. J. The psychology of verbal comprehension. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 3). Hillsdale, NJ: Erlbaum.
Sternberg, R. J. Review of Nonverbal Test of Cognitive Skills. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook. Lincoln: University of Nebraska Press, in press.
Sternberg, R. J. Review of Test of Cognitive Skills. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook. Lincoln: University of Nebraska Press, in press.
Sternberg, R. J. Teaching critical thinking: Are we making critical mistakes? Phi Delta Kappan, in press.
Sternberg, R. J. Teaching intelligence: The application of cognitive psychology to the improvement of intellectual skills. In J. B. Baron & R. J. Sternberg (Eds.), Critical thinking: Its nature, assessment, and improvement. New York: W. H. Freeman, in press.
Sternberg, R. J. A triarchic view of intelligence in cross-cultural perspective. In S. Irvine (Ed.), The cultural context of human abilities. New York: Wiley, in press.
Sternberg, R. J. A unified theory of intellectual exceptionality. In J. G. Borkowski & J. Day (Eds.), Cognition and intelligence in special children: Comparative approaches to retardation, learning disabilities, and giftedness. Norwood, NJ: Ablex, in press.
Sternberg, R. J. What is adaptive? Behavioral and Brain Sciences, in press.
Sternberg, R. J. Yes, no, you must be kidding: Three responses to three critiques. Behavioral and Brain Sciences, in press.
Sternberg, R. J., & Baron, J. B. A triarchic approach to measuring critical thinking skills: A psychological view. Educational Leadership, in press.
Sternberg, R. J., & Berg, C. A. Definitions of intelligence: A quantitative comparison of the 1921 and 1986 symposia. In R. J. Sternberg & D. K. Detterman (Eds.), What is intelligence? Contemporary viewpoints on its nature and definition. Norwood, NJ: Ablex, in press.

Sternberg, R. J., & Berg, C. A. What are theories of adult intellectual development theories of? In C. Schooler & K. W. Schaie (Eds.), Intellectual functioning, social structure, and aging. Washington, D.C.: National Institutes of Health.

Sternberg, R. J., & Davidson, J. E. (Eds.), Conceptions of giftedness. New York: Cambridge University Press, in press.

Sternberg, R. J., & Davidson, J. E. Conceptions of giftedness: A map of the terrain. In R. J. Sternberg & J. E. Davidson (Eds.), Conceptions of giftedness. New York: Cambridge University Press, in press.

Sternberg, R. J., & Detterman, D. K. (Eds.). What is intelligence? Contemporary viewpoints on its nature and definition. Norwood, NJ: Ablex, in press.

Sternberg, R. J., & Smith, E. E. (Eds.), The psychology of human thought. New York: Cambridge University Press, in press.

Sternberg, R. J., & Suben, J. The socialization of intelligence. In M. Perlmutter (Ed.), Minnesota symposium on child development. New York: Academic Press, in press.

Sternberg, R. J., & Wagner, R. K. (Eds.). Practical intelligence: Its nature, origins, and assessment. New York: Cambridge University Press, in press.

Wagner, R. K., & Sternberg, R. J. Executive processes in reading. In B. Britton (Ed.), Executive control processes in reading. Hillsdale, NJ: Erlbaum, in press.

Dr. Phillip L. Ackerman
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

AFOSR,
Life Sciences Directorate
Bolling Air Force Base
Washington, DC 20332

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

Dr. John R. Anderson
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. Nancy S. Anderson
Department of Psychology
University of Maryland
College Park, MD 20742

Dr. Phipps Arabie
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Special Assistant for Projects,
OASN(M&RA)
5D800, The Pentagon
Washington, DC 20350

Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

Dr. Patricia Baggett
University of Colorado
Department of Psychology
Box 345
Boulder, CO 80309

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

Dr. John Black
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Arthur S. Blaiwes
Code N711
Naval Training Equipment Center
Orlando, FL 32813

Dr. R. Darrell Bock
University of Chicago
Department of Education
Chicago, IL 60637

Dr. Gordon H. Bower
Department of Psychology
Stanford University
Stanford, CA 94306

Dr. Robert Breaux
Code N-095R
NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Ann Brown
Center for the Study of Reading
University of Illinois
51 Gerty Drive
Champaign, IL 61280

Dr. John S. Brown
XEROX Palo Alto Research
Center
333 Coyote Road
Palo Alto, CA 94304

Dr. Bruce Buchanan
Computer Science Department
Stanford University
Stanford, CA 94305

Dr. Robert D. Bubr
Hubbs Sea World Institute
1700 S. Shores Rd.
Mission Bay
San Diego, CA 92109

Dr. Patricia A. Butler
NIE Mail Stop 1806
1200 19th St., NW
Washington, DC 20208

Dr. Tom Cafferty
Dept. of Psychology
University of South Carolina
Columbia, SC 29208

Dr. Robert Calfee
School of Education
Stanford University
Stanford, CA 94305

Dr. Susan Carey
Harvard Graduate School of
Education
337 Gutman Library
Appian Way
Cambridge, MA 02138

Dr. Pat Carpenter
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. Robert Carroll
NAVOP 01B7
Washington, DC 20370

Dr. Fred Chang
Navy Personnel R&D Center
Code 51
San Diego, CA 92152

Dr. Davida Charney
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Michelene Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Professor Chu Tien-Chen
Mathematics Department
National Taiwan University
Taipei, TAIWAN

Dr. Yee-Yeen Chu
Perceptronics, Inc.
21111 Erwin Street
Woodland Hills, CA 91367-3713

Dr. William Clancey
Computer Science Department
Stanford University
Stanford, CA 94306

Chief of Naval Education
and Training
Liaison Office
Air Force Human Resource Laboratory
Operations Training Division
Williams AFB, AZ 85224

Dr. Michael Cole
University of California
at San Diego
Laboratory of Comparative
Human Cognition - D003A
La Jolla, CA 92093

Dr. Michael Coles
University of Illinois
Department of Psychology
Champaign, IL 61820

Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Leon Cooper
Brown University
Center for Neural Science
Providence, RI 02912

Dr. Lynn A. Cooper
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Capt. Jorge Correia Jesuino
Marinha-7A Reparticao
Direccao Do Servico Do Pessoal
Praca Do Comercio
Lisbon
PORTUGAL

Dr. Meredith P. Crawford
American Psychological Association
Office of Educational Affairs
1200 17th Street, N.W.
Washington, DC 20036

M.C.S. Louis Crocq
Secretariat General de la
Defense Nationale
51 Boulevard de Latour-Maubourg
75007 Paris
FRANCE

Dr. Lee Cronbach
16 Laburnum Road
Atherton, CA 94205

CAPT P. Michael Curran
Office of Naval Research
800 N. Quincy St.
Code 125
Arlington, VA 22217-5000

Bryan Dallman
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Sharon Derry
Florida State University
Department of Psychology
Tallahassee, FL 32306

Dr. R. K. Dismukes
Associate Director for Life Sciences
AFOSR
Bolling AFB
Washington, DC 20332

Dr. Emanuel Donchin
University of Illinois
Department of Psychology
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Thomas M. Duffy
Communications Design Center
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Richard Duran
University of California
Santa Barbara, CA 93106

Dr. John Ellis
Navy Personnel R&D Center
San Diego, CA 92252

Dr. Jeffrey Elman
University of California,
San Diego
Department of Linguistics, C-008
La Jolla, CA 92093

Dr. Richard Elster
Deputy Assistant Secretary
of the Navy (Manpower)
OASN (M&RA)
Department of the Navy
Washington, DC 20350-1000

Dr. Susan Embretson
University of Kansas
Psychology Department
Lawrence, KS 66045

Dr. Randy Engle
Department of Psychology
University of South Carolina
Columbia, SC 29208

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. K. Anders Ericsson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. W. E. Evans
Hubbs Sea World Institute
1700 S. Shores Rd.
Mission Bay
San Diego, CA 92109

Dr. Martha Farah
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Beatrice J. Farr
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Marshall J. Farr
2520 North Vernon Street
Arlington, VA 22207

Dr. Pat Federico
Code 511
NPRDC
San Diego, CA 92152

Mr. Wallace Feurzeig
Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02238

Dr. Gerhard Fischer
University of Colorado
Department of Computer Science
Boulder, CO 80309

Dr. Baruch Fischhoff
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91367

Dr. Dexter Fletcher
University of Oregon
Computer Science Department
Eugene, OR 97403

Dr. Linda Flower
Carnegie-Mellon University
Department of English
Pittsburgh, PA 15213

Dr. Carl H. Frederiksen
McGill University
3700 McTavish Street
Montreal, Quebec H3A 1Y2
CANADA

Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

Dr. Norman Frederiksen
Educational Testing Service
Princeton, NJ 08541

Dr. R. Edward Geiselman
Department of Psychology
University of California
Los Angeles, CA 90024

Dr. Michael Genesereth
Stanford University
Computer Science Department
Stanford, CA 94305

Dr. Dedre Gentner
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Don Gentner
Center for Human
Information Processing
University of California
La Jolla, CA 92093

Dr. Robert D. Gibbons
University of Illinois-Chicago
P.O. Box 6998
Chicago, IL 60680

Dr. Herbert Ginsburg
University of Rochester
Graduate School of
Education
Rochester, NY 14627

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Arthur M. Glenberg
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

Dr. Marvin D. Glock
13 Stone Hall
Cornell University
Ithaca, NY 14853

Dr. Gene L. Gloye
Office of Naval Research
Detachment
1030 E. Green Street
Pasadena, CA 91106-2485

Dr. Sam Glucksberg
Princeton University
Department of Psychology
Green Hall
Princeton, NJ 08540

Dr. Joseph Goguen
Computer Science Laboratory
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. Susan Goldman
University of California
Santa Barbara, CA 93106

Dr. Daniel Gopher
Industrial Engineering
& Management
TECHNION
Haifa 32000
ISRAEL

Dr. Sherrie Gott
AFHRL/MODJ
Brooks AFB, TX 78235

Dr. T. Govindaraj
Georgia Institute of Technology
School of Industrial & Systems
Engineering
Atlanta, GA 30332

Dr. Richard H. Granger
Department of Computer Science
University of California, Irvine
Irvine, CA 92717

Dr. Wayne Gray
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. James G. Greeno
University of California
Berkeley, CA 94720

Dr. Henry M. Halff
Halff Resources, Inc.
4918 33rd Road, North
Arlington, VA 22207

Dr. Cheryl Hamel
NTSC
Orlando, FL 32813

NR 154-505 [Yale/RJ Sternberg] Distribution List

Dr. Ray Hannapel
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Stevan Harnad
Editor, The Behavioral and
Brain Sciences
20 Nassau Street, Suite 240
Princeton, NJ 08540

Janice Hart
Office of the Chief
of Naval Operations
OP-11HD
Department of the Navy
Washington, D.C. 20350-2000

Dr. Reid Hastie
Northwestern University
Department of Psychology
Evanston, IL 60201

Prof. John R. Hayes
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Joan I. Heller
505 Haddon Road
Oakland, CA 94606

Dr. Jim Hollan
Intelligent Systems Group
Institute for
Cognitive Science (C-015)
UCSD
La Jolla, CA 92093

Dr. Melissa Holland
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Keith Holyoak
University of Michigan
Human Performance Center
330 Packard Road
Ann Arbor, MI 48109

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Earl Hunt
Department of Psychology
University of Washington
Seattle, WA 98105

Dr. Ed Hutchins
Intelligent Systems Group
Institute for
Cognitive Science (C-015)
UCSD
La Jolla, CA 92093

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Dillon Inouye
WICAT Education Institute
Provo, UT 84057

Dr. Alice Isen
Department of Psychology
University of Maryland
Catonsville, MD 21228

Dr. Zachary Jacobson
Bureau of Management Consulting
365 Laurier Avenue West
Ottawa, Ontario K1A 0S5
CANADA

Dr. Robert Jannarone
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. Claude Janvier
Directeur, CIRADE
Universite' du Quebec a Montreal
Montreal, Quebec H3C 3P8
CANADA

NR 154-505 [Yale/RJ Sternberg] Distribution List

Dr. Robin Jeffries
Computer Research Center
Hewlett-Packard Laboratories
1501 Page Mill Road
Palo Alto, CA 94304

Margaret Jerome
c/o Dr. Peter Chandler
83, The Drive
Hove
Sussex
UNITED KINGDOM

Dr. Douglas H. Jones
Advanced Statistical
Technologies Corporation
10 Trafalgar Court
Lawrenceville, NJ 08148

Dr. Marcel Just
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Daniel Kahneman
The University of British Columbia
Department of Psychology
#154-2053 Main Mall
Vancouver, British Columbia
CANADA V6T 1Y7

Dr. Milton S. Katz
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Steven W. Keele
Department of Psychology
University of Oregon
Eugene, OR 97403

Dr. Wendy Kellogg
IBM T. J. Watson Research Ctr.
P.O. Box 218
Yorktown Heights, NY 10598

Dr. Scott Kelso
Haskins Laboratories,
270 Crown Street
New Haven, CT 06510

Dr. Dennis Kibler
University of California
Department of Information
and Computer Science
Irvine, CA 92717

Dr. David Kieras
University of Michigan
Technical Communication
College of Engineering
1223 E. Engineering Building
Ann Arbor, MI 48109

Dr. Peter Kincaid
Training Analysis
& Evaluation Group
Department of the Navy
Orlando, FL 32813

Dr. Walter Kintsch
Department of Psychology
University of Colorado
Campus Box 345
Boulder, CO 80302

Dr. David Klahr
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Mazie Knerr
Program Manager
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. Janet L. Kolodner
Georgia Institute of Technology
School of Information
& Computer Science
Atlanta, GA 30332

Dr. Stephen Kosslyn
Harvard University
1236 William James Hall
33 Kirkland St.
Cambridge, MA 02138

Dr. Kenneth Kotovsky
Department of Psychology
Community College of
Allegheny County
800 Allegheny Avenue
Pittsburgh, PA 15233

Dr. David H. Krantz
2 Washington Square Village
Apt. # 15J
New York, NY 10012

Dr. Gerald Kulm
NIE
Mail Stop 1806
1200 19th Street, NW
Washington, DC 20208

Dr. David R. Lambert
Naval Ocean Systems Center
Code 441T
271 Catalina Boulevard
San Diego, CA 92152

Dr. Pat Langley
University of California
Department of Information
and Computer Science
Irvine, CA 92717

Dr. Marcy Lansman
University of North Carolina
The L. L. Thurstone Lab.
Davie Hall 013A
Chapel Hill, NC 27514

Dr. Jill Larkin
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. Jean Lave
School of Social Sciences
University of California
Irvine, CA 92717

Dr. Robert Lawler
Information Sciences, FRL
GTE Laboratories, Inc.
40 Sylvan Road
Waltham, MA 02254

Dr. Alan M. Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Jim Levin
University of California
Laboratory for Comparative
Human Cognition
D003A
La Jolla, CA 92093

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Clayton Lewis
University of Colorado
Department of Computer Science
Campus Box 430
Boulder, CO 80309

Science and Technology Division
Library of Congress
Washington, DC 20540

Dr. Marcia C. Linn
Lawrence Hall of Science
University of California
Berkeley, CA 94720

Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. Don Lyon
P. O. Box 44
Higley, AZ 85236

Lt. (Army) Jesus Magro Morales
C/Roncesvalles, 5, 4 - 1
28007 Madrid
SPAIN

Dr. William L. Maloy
Chief of Naval Education
and Training
Naval Air Station
Pensacola, FL 32508

Dr. Sandra P. Marshall
Dept. of Psychology
San Diego State University
San Diego, CA 92182

Dr. Jay McClelland
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. James L. McGaugh
Center for the Neurobiology
of Learning and Memory
University of California, Irvine
Irvine, CA 92717

Dr. Joe McLachlan
Navy Personnel R&D Center
San Diego, CA 92152

Dr. James McMichael
Assistant for MPT Research,
Development, and Studies
NAVOP 01B7
Washington, DC 20370

Dr. Barbara Means
Human Resources
Research Organization
1100 South Washington
Alexandria, VA 22314

Dr. Jose Mestre
University of Massachusetts
401 Goodell Building
Amherst, MA

Prof. D. Michie
The Turing Institute
36 North Hanover Street
Glasgow G1 2AD, Scotland
UNITED KINGDOM

Dr. George A. Miller
Department of Psychology
Green Hall
Princeton University
Princeton, NJ 08540

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Dr. Andrew R. Molnar
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152

Dr. Tom Moran
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

Dr. Allen Munro
Behavioral Technology
Laboratories - USC
1845 S. Elena Ave., 4th Floor
Redondo Beach, CA 90277

Dr. David Navon
Institute for Cognitive Science
University of California
La Jolla, CA 92093

Assistant for MPT Research,
Development and Studies
NAVOP 01B7
Washington, DC 20370

Dr. Allen Newell
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. T. Niblett
The Turing Institute
36 North Hanover Street
Glasgow G1 2AD, Scotland
UNITED KINGDOM

Dr. Richard E. Nisbett
University of Michigan
Institute for Social Research
Room 5261
Ann Arbor, MI 48109

Dr. Donald A. Norman
Institute for Cognitive Science
University of California
La Jolla, CA 92093

Dr. Melvin R. Novick
356 Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Director, Training Laboratory,
NPRDC (Code 05)
San Diego, CA 92152

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152

Director, Human Factors
& Organizational Systems Lab,
NPRDC (Code 07)
San Diego, CA 92152

Fleet Support Office,
NPRDC (Code 301)
San Diego, CA 92152

Library, NPRDC
Code P201L
San Diego, CA 92152

Commanding Officer,
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. Harry F. O'Neil, Jr.
University of Southern California
School of Education -- WPH 801
Dept. of Educational
Psychology and Technology
Los Angeles, CA 90089-0031

Dr. Stellan Ohlsson
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Office of Naval Research,
Code 1142
800 N. Quincy St.
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142EP
800 N. Quincy Street
Arlington, VA 22217-5000

Group Psychology Program,
ONR Code 1142GP
800 N. Quincy St.
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142PT
800 N. Quincy
Arlington, VA 22217-5000
(6 Copies)

Office of Naval Research,
Code 125
800 N. Quincy Street
Arlington, VA 22217-5000

Psychologist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

Special Assistant for Marine
Corps Matters,
ONR Code 00MC
800 N. Quincy St.
Arlington, VA 22217-5000

Psychologist
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Lt. Col. (Dr.) David Payne
AFHRL
Brooks AFB, TX 78235

Dr. James W. Pellegrino
University of California,
Santa Barbara
Department of Psychology
Santa Barbara, CA 93106

Dr. Robert Penn
NPRDC
San Diego, CA 92152

Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637

Dr. Ray Perez
ARI (PERI-II)
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. David N. Perkins
Educational Technology Center
337 Gutman Library
Appian Way
Cambridge, MA 02138

Dr. Nancy Perry
Chief of Naval Education
and Training, Code 00A2A
Naval Station Pensacola
Pensacola, FL 32508

Dr. Martha Polson
Department of Psychology
Campus Box 346
University of Colorado
Boulder, CO 80309

Dr. Peter Polson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Mike Posner
University of Oregon
Department of Psychology
Eugene, OR 97403

Dr. Sukai Prom-Jackson
1821 Massachusetts Ave., NW
#602
Washington, DC 20005

Dr. Joseph Psotka
ATTN: PERI-1C
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Fred Reif
Physics Department
University of California
Berkeley, CA 94720

Dr. Lauren Resnick
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. Mary S. Riley
Program in Cognitive Science
Center for Human Information
Processing
University of California
La Jolla, CA 92093

William Rizzo
Code 712 NAVTRAEQUIPCEN
Orlando, FL 32813

Dr. Andrew M. Rose
American Institutes
for Research
1055 Thomas Jefferson St., NW
Washington, DC 20007

Dr. William B. Rouse
Georgia Institute of Technology
School of Industrial & Systems
Engineering
Atlanta, GA 30332

Dr. David Rumelhart
Center for Human
Information Processing
Univ. of California
La Jolla, CA 92093

Dr. E. L. Saltzman
Haskins Laboratories
270 Crown Street
New Haven, CT 06510

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
Knoxville, TN 37916

Dr. Michael J. Samet
Perceptronics, Inc
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Robert Sasmor
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Roger Schank
Yale University
Computer Science Department
P.O. Box 2158
New Haven, CT 06520

Frank L. Schmidt
Department of Psychology
Bldg. GG
George Washington University
Washington, DC 20052

Dr. Walter Schneider
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Alan H. Schoenfeld
University of California
Department of Education
Berkeley, CA 94720

Dr. Janet Schofield
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Judah L. Schwartz
MIT
20C-120
Cambridge, MA 02139

Dr. Marc Sebrechts
Department of Psychology
Wesleyan University
Middletown, CT 06475

Dr. Judith Segal
Room 819F
NIE
1200 19th Street N.W.
Washington, DC 20208

Dr. Ramsay W. Selden
Assessment Center
CCSSO
Suite 379
400 N. Capitol, NW
Washington, DC 20001

Dr. W. Steve Sellman
OASD(MRA&L)
2B269 The Pentagon
Washington, DC 20301

Dr. Sylvia A. S. Shafto
National Institute of Education
1200 19th Street
Mail Stop 1806
Washington, DC 20208

Dr. T. B. Sheridan
Dept. of Mechanical Engineering
MIT
Cambridge, MA 02139

Dr. Lee Shulman
Stanford University
1040 Cathcart Way
Stanford, CA 94305

Dr. Randall Shumaker
Naval Research Laboratory
Code 7510
4555 Overlook Avenue, S.W.
Washington, DC 20375-5000

Dr. Miriam Schustack
Code 51
Navy Personnel R & D Center
San Diego, CA 92152

Dr. Robert S. Siegler
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Edward Silver
Dept. of Mathematics
San Diego State University
San Diego, CA 92115

Dr. Zita M Simutis
Instructional Technology
Systems Area
ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Derek Sleeman
Stanford University
School of Education
Stanford, CA 94305

Dr. Edward E. Smith
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Richard E. Snow
Department of Psychology
Stanford University
Stanford, CA 94306

Dr. Elliot Soloway
Yale University
Computer Science Department
P.O. Box 2158
New Haven, CT 06520

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Paul Speckman
University of Missouri
Department of Statistics
Columbia, MO 65201

Dr. Kathryn T. Spoehr
Brown University
Department of Psychology
Providence, RI 02912

James J. Staszewski
Research Associate
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Saul Sternberg
University of Pennsylvania
Department of Psychology
3815 Walnut Street
Philadelphia, PA 19104

Dr. Kurt Steuck
AFHRL/MOD
Brooks AFB
San Antonio TX 78235

NR 154-505 [Yale/RJ Sternberg] Distribution List

Dr. Paul J. Sticha
Senior Staff Scientist
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Cdr Michael Suman, PD 303
Naval Training Equipment Center
Code N51, Comptroller
Orlando, FL 32813

Dr. Patrick Suppes
Stanford University
Institute for Mathematical
Studies in the Social Sciences
Stanford, CA 94305

Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152

Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research
Laboratory
Urbana, IL 61801

Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

Dr. Douglas Towne
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

Dr. Robert Tsutakawa
The Fred Hutchinson
Cancer Research Center
Division of Public Health Sci.
1124 Columbia Street
Seattle, WA 98104

Dr. James Tweeddale
Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Paul Twohig
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 310
St. Paul, MN 55114

Dr. Kurt Van Lehn
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

Dr. Ming-Mei Wang
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Beth Warren
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Shih-Sung Wen
Jackson State University
1325 J. R. Lynch Street
Jackson, MS 39217

Dr. Barbara White
Bolt Beranek & Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Champaign, IL 61820

NR 154-505 [Yale/RJ Sternberg] Distribution List

Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152

Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

Dr. Merlin C. Wittrock
Graduate School of Education
UCLA
Los Angeles, CA 90024

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152

Dr. Joe Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Joseph L. Young
Memory & Cognitive
Processes
National Science Foundation
Washington, DC 20550

Dr. Steven Zornetzer
Office of Naval Research
Code 1140
800 N. Quincy St.

END

FILMED

2-86

DTIC